THE IMPORTANCE OF SELECTING THE OPTIMAL NUMBER OF PRINCIPAL COMPONENTS FOR FAULT DETECTION USING PRINCIPAL COMPONENT ANALYSIS

PATRICIA HELEN KHWAMBALA

Thesis presented in fulfillment of the requirements for the degree of Master of Science in Electrical Engineering at the University of Cape Town.

Supervisor: Prof. M. Braae

Department of Electrical and Electronics Engineering

University Of Cape Town

DECLARATION

I know the meaning of plagiarism and declare that the work presented in this thesis is my original work, save for that which has been properly acknowledged.

SIGNATURE:…….

DATE:……..

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to the Divine Consciousness that lives in all things and loves in all ways, for the opportunity.

I also express my appreciation to my supervisor Professor Martin Braae for the guidance to come up with this.

To my spiritual and physical family and friends for believing in me, without your support this could not be possible, I will always be grateful.

DEDICATION

To my late **Mum,** you will always be my pillar, my daughter **Maziko Joan Kagoli** and my son **Mzati Kagoli** – *I love you all*.

SUMMARY

Fault detection and isolation are the two fundamental building blocks of process monitoring. Accurate and efficient process monitoring increases plant availability and utilization.

Principal component analysis is one of the statistical techniques that are used for fault detection. Determination of the number of PCs to be retained plays a big role in detecting a fault using the PCA technique. In this dissertation focus has been drawn on the methods of determining the number of PCs to be retained for accurate and effective fault detection in a laboratory thermal system.

SNR method of determining number of PCs, which is a relatively recent method, has been compared to two commonly used methods for the same, the CPV and the scree test methods. SNR method gives the number of PCs that are more sensitive to a particular fault in a system. SNR method gives different number of PCs to be retained for different types of faults in the same system. Through this method the efficiency of the fault detection using PCA is achieved, especially when the very same types of faults are experienced.

The detection of the fault can be calculated by the use of T- squared statistics and the squared prediction error (Q- statistic) techniques. In this dissertation T – squared statistics has been used for fault detection.

ABBREVIATIONS

AEM	Abnormal Event Management
ARX	Autoregressive with Exogenous input
ASYM	Asymptotic Model
BLFRLS	Bi-Loop Forgetting Factor Recursive Least Squares
EKF	Extended Kalman Filter
FD	Fault Detection
FDI	Fault Detection and Isolation
FI	Fault Isolation
LS	Least Square
LTI	Linear and Time Invariant
MIMO	Multiple Input, Multiple Output
MOESP	Multivariable Output Error State Space
N4SID	Numerical Algorithms for subspace system identification
MPC	Model Predictive Control
ORT	Orthogonal Decomposition
PEM	Prediction Error Method
PCA	Principal Component Analysis
PCs	Principal Components
PLS	Partial Least Squares
RLS	Recursive Least Square
SID	System Identification
SISO	Single Input, Single Output
SVD	Singular value Decomposition
TSODS	Two Stage Orthogonal Decomposition Subspace
UKF	Unscented Kalman Filter

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

The dissertation is about the fault detection in the laboratory thermal system using the principal component analysis. The three methods of determining the optimal number of principal components to be retained when this method is used for fault detection have been experimentally compared.

Fault detection – This is the detection of any interruption in a system ability to perform a required task. This is one of the important parts of fault management for all process control systems. An effective fault detection method should be among other things reliable, reconfigurable and sensitive to a fault [102] for a quick detection of a fault in a system. When a fault is detected, fault identification and diagnosis follows as described in the next chapter.

Fault detection in a process system cannot be ignored because it helps to protect the life of the system and that of the people operating the machine. Above that the short time taken in detecting and diagnosing a fault in a system can help to save time and money, which in return increase the profitability of the machine [100].

There are a number of classes of failure as outlined in chapter 2, which is the total interruption of system ability to perform a required task under specified operating conditions [101] which leads to a fault.

In this dissertation, fault detection using principal component analysis has been applied on the laboratory thermal system. This would help it make the equipment always ready for researchers with reliable results.

There are a number of methods that has been used in process control systems for fault detection. These methods of fault detection are either model based or model – free based methods [51], some of these methods are as outlined in chapter 2. The focus in this dissertation is the use of principal component analysis in fault detection, because of its quick ability to detect sensor faults, which makes part of the thermal system.

Principal Component analysis - This is statistical technique which is mainly used for dimension reduction [1, 9]. Apart from fault detection, PCA can also be used for system identification and fault diagnosis. The technique that is used in this dissertation and its application is on the laboratory thermal system data.

In fault detection PCA makes use of the Q - statistics and T^2 - statistics. Q statistics uses the loading vectors that are associated with the smaller singular values [5, 39 and 41]. While the T^2 statistics technique uses the loading vectors associated with the largest singular values [42].

There are a number of methods used to determine the number of PCs to be used for fault detection when using PCA [9].

Some of the methods commonly used for determining the number of PCs are Scree plot, Average Eigenvalue (AE), Fault Signal to Noise Ratio (fault SNR) and Cumulative Percent Variance (CPV) methods just to mention a few [39 and 79]. Fault SNR is relatively the most recent method of determining the number of PCs for fault detection. ***The aim of this dissertation is to investigate on the effect of determining the optimal number of principal components to be retained for fault detection on a laboratory thermal system***. A number of researchers have done the same before [12, 39, and 79] but nothing has been done on the university of Cape Town control laboratory thermal system hence the dissertation.

This dissertation is divided into seven chapters, **introduction** in chapter 1. Chapter 2 gives the outline of **fault detection** in process systems, including the relevant literature review. The literature review is in chapter 3 analyzing the **principal component analysis** in fault detection on a thermal system and things to be considered.

Chapter 4 gives the overview of the **laboratory thermal system** which is used in this dissertation and its current status and the system stability. In chapter 5 **PCA fault detection technique is applied on the laboratory thermal system process data.** The focus is on sensitivity to a particular fault with different number of PCs retained. **The comparison on different methods used** for determining the number of PCs to be retained for fault detection is done in chapter 6. The **discussions on the results are** laid out in chapter 7, followed by **recommendations.**

CHAPTER 2

SYSTEM FAULT DETECTION

The thermal laboratory equipment is to be used for the application of PCA for fault detection while the machine is in operation. This is good for researchers, because results from this system will be correct. The fault detection method will help to detect any fault that occurs when the system is in use, so that every data obtained reflects the real state of the system.

There are a number of methods that have been used for fault detection in thermal systems alongside principal components analysis. Experiments performed in [96] used the operator based approach to detect faults. The experimental results from the thermal system confirmed that the operator based robust RCF factorization approach is effective for this purpose.

Adaptive observer based fault detection and diagnosis approach; this is an online model based procedure for fault detection of thermal process faults. The autoregressive with exogenous input – ARX models are used as dynamic features for fault detection [106].

The principal component analysis has also been used for fault detection in a thermal system in [97]. The experimental results indicated that this method can effectively capture the non linear relationship among the process variables hence effective for fault detection.

Parity Space: This method checks the parity of the mathematical equations of the system (analytical redundancy relations), by using the actual measurements. This method is robust but with a computational burden [105].

State observers: In this method outputs of the system are reconstructed from the measurements with the aid of the observers or estimators. This method has a low computational burden but cannot detect sensor failures reliably [104].

Adaptive Parity Space: This method uses the parity space approach as a temporal analytical redundancy. The formulation and calculation of the parity spaces are done in either the continuous or discrete time domains. It detects faults slowly but able to detect multiple simultaneous faults. The disadvantage of this method is that it is insensitive to noise and complex to implement [105].

Fault Detection Method and Abilities Table

Fault Detection Method	Abilities
Adaptive Observer	Unable to detect multiple simultaneous faults. Able to detect faults promptly. Has a good isolability and average sensitivity. Can be easily implemented.
Principal Component Analysis	Does not facilitate isolation of the fault. Able to detect actuator, process and sensor faults fairly quickly. Implementation is complex.
Parity Space	Flexible in design of its structure, which accounts for its ability to detect faults accurately but spurious faults. Has high sensitivity towards noise. Unable to detect multiple simultaneous faults. Implementation is complex when detecting multiple sensor and actuator faults. Has computational burden. Is robust and has an average isolability.
State Observers	Robust to parameter variations and noise Cannot detect sensor failures reliably. Implementation becomes complex when detecting multiple faults. Not very sensitive to fault effects and modelling uncertainties. Low computational burden and can easily be implemented. Unable to detect multiple simultaneous faults. Prompt detection with bad to average isolability.
Adaptive Parity Space	Detects faults slowly. Able to detect multiple simultaneous faults. Insensitive to noise and has good isolability. Implementation is complex

Comment	All the various methods used for fault detection have their advantages and disadvantages as given in the table above. The Adaptive parity space and state observers which are slow to detect faults and cannot detect sensor failures effectively despite being easy to implement. This is why Principal component analysis method was used for fault detection in the thermal system because the sensors. This method is able to detect sensor faults despite being complex to implement.

2.1 Desirable Properties of Fault Detection

In 1988 Gertler [102] and B. Patel in 2001 [100] described the following as the desirable properties of fault detection:

Sensitivity – Any fault detection method should be sensitive to faults that occur in the system. This is a qualitative measure that makes the size of the fault to be known. Sensitivity depends on the size of the elements in the system matrices, and noise in the system.

Reliability – This is the most important quality of fault detection, being reliable at all times to detect any occurring fault. Good reliability enables a normal completion of the tasks after failure, so the ability of the method to complete the given task perfectly over a period of time defines its reliability.

Reconfigurable – The parameters of the fault detecting method must adjust to faults. If the method is able to do that it is called both adaptive and redundant. When the method is redundant, it is able to overcome the lost capabilities with its remaining resources and if it is adaptive, it can adjust to out- of-normal behaviour.

Isolability – The fault detection method should be able to isolate specific faults. This ability to isolate a fault depends on the structure of a plant and the statistical methods for fault detection. It is important to have a threshold used for fault detection though it is difficult to come up with one. Having a low threshold means only intermediate and large faults will be detected and small faults can not be picked. Setting a high threshold means only large faults will be detected while small and intermediate faults will not be detected.

Survivability – This gives the ability of the fault detection method to operate safely. To be safe to the personnel operating the machine and also safe to the system itself. This should be irrespective of whether

the task has been completed or not. The good method should allow the degradation in performance after a fault has occurred. But the system should be in a safe state of operation.

Robustness - In the presence of component failures and modelling, fault method should still be able to detect faults. The inevitability of modelling errors can interfere with fault detection.

2.2 Classes of Failures

R. Iserman (2005) [101] described a failure as a total interruption of system ability to perform a required task under specified operating conditions. This is an event and usually a result from one or many faults. Gertler (1988) and B.Patel (2001) [100 1nd 102] also outlined the main three classes of faults:

1. Additive process faults – These are defined as the disturbances acting on the plant. Additive process faults normally do not exist and if they exist causes a shift in the plant independent of the measured inputs. They are sometimes referred to as leaked faults.

2. Additive measurements faults – These are the differences that arise between the true and the measured values of the input variables or the plant output. Actuator malfunction description can be one of the uses of this class of faults.

3. Multiplicative Process faults – Plant parameters can be gradually or abruptly be changed which can cause this type of faults. Surface contamination and total loss of power perfectly describes such type of faults.

Any system is divided into three subsystems: actuators, actual process and sensors, which is sometimes called instrumentation. The process faults is the most difficult to compensate while the sensor fault can easily be corrected by electronic switching techniques, this can not involve the reconfiguration of the mechanical parts. There is no simple compensation of the actuator type faults; this is more difficult than a redirection of electric signals.

2.3 Process Fault Detection and Management

Technical processes need to be supervised to show its present state, showing desirable or undesirable states, so that appropriate action can be taken. This will help to avoid accidents to the operating personnel and damage to the machine. A fault first affects the internal process parameters and causes changes in resistance, capacitance and internal variables [101] These affect the measurable output, by that small change.

2.3.1 Fault Detection in System Processes

There are two main classes of methods for fault detection and identification, namely model - free methods and model based methods.

Model – Free Methods of fault detection identification do not use mathematical model of the plant. They range from hardware physical redundancy through signal processing methods. While Model – Based Methods use a mathematical model of a monitored plant and rely on the concept of analytically computed values of the respective variable [51].

Model Approach - Model based is one of the three methods of fault detection of which the others are knowledge based and history based methods [101]. Model based method of fault detection developments began in the early 1970's [102]. This method uses mathematical model of the plant under monitoring. Mathematical model of the monitored system can be obtained along two routes or a combination of both. The one route of obtaining a mathematical model is to divide the system into important subsystems, in the subsystems the properties are well understood from physical laws and natural laws. Then these subsystems can be combined mathematically to obtain one complete model of the whole system.

Experimentation is the second route of obtaining a mathematical model. Input and output signals of the system are obtained and data analysis is done to derive a model [107]. The figure 2.1 below shows this concept of model based fault detection. The details are explained in appendix B. The important thing to note is the use of the model within the fault detection system.

Model based method of fault detection can further be divided as either qualitative models and quantitative models [108].

Differential equations, state space methods, transfer functions etc of quantitative models are used to generally utilize results from the field of control theory [106].These models use static and dynamic relations among system variables and parameters in order to describe the systems behavior in quantitative mathematical terms. While in qualitative models, the realization between variables to obtain the expected system behavior is expressed in terms of qualitative functions centered on different units in the process like abstraction hierarchy and causal models. Taking into consideration that large systems are highly non – linear and dynamics present in the process under monitoring, qualitative models are usually used.

There are a number of advantages upon the use of explicit models in fault detection methods [109] as follows;

Higher fault detection performance can be obtained, where mode types of faults can be detected with shorter detection times.

Fault detection can be performed over a large operating range and fault detection can be performed passively without disturbing the process.

There is also a disadvantage of model – based fault detection methods in that the preliquisite of accurate models of the process being monitored and possibly more complex design procedure [103]. The major performance limiting factor for model based methods is the model accuracy.

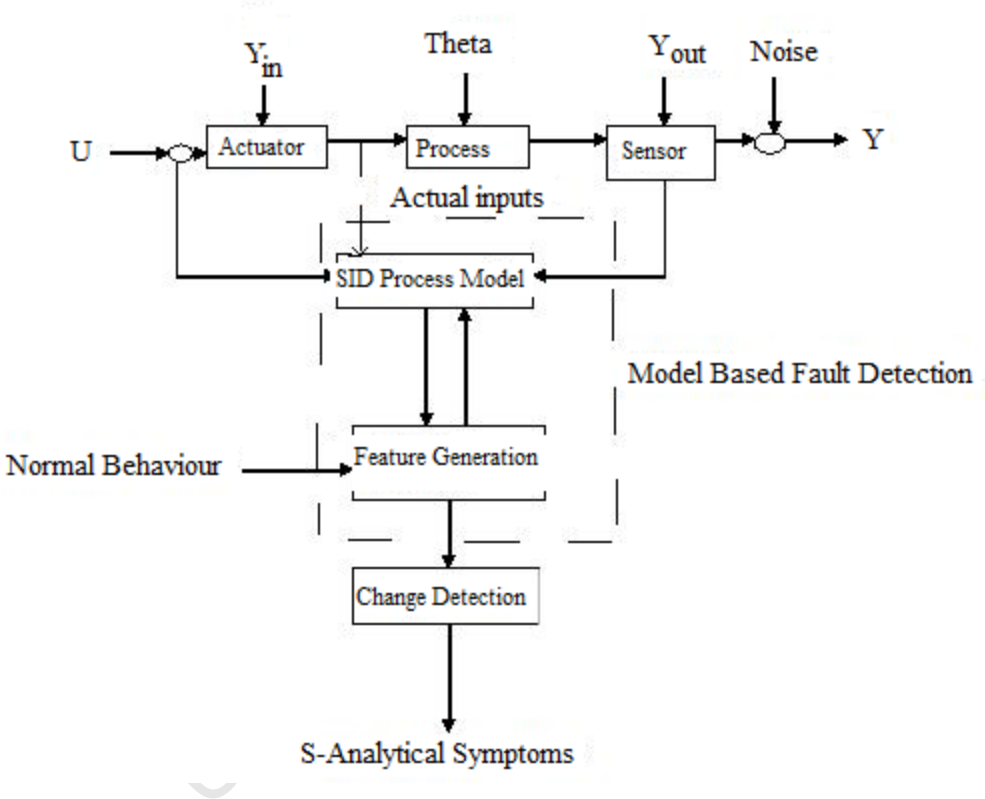

Figure 2.1: Model Based fault Detection Structure [56].

2.3.2 Fault Management

The supervision of technical processes is done by limit-checking or threshold-checking of the measurable outputs. One checks if the quantities are within tolerance or not. If quantities go beyond the tolerance then a fault occurred. There are a number of stages that are involved in this act of process supervision:

Monitoring – An alarm is set to alert the operator, which is triggered when the threshold is exceeded.

Automatic Protection – This comes in when the exceeding of the threshold leads to a dangerous process state. The monitoring function automatically starts an appropriate counteraction which is usually an immediate shut down.

Supervision with fault diagnosis – This involves feature generation, fault detection, generation of symptoms that lead to fault diagnosis which goal is to determine the kind, size and location of the fault. Fault evaluation follows that helps to come up with a decision on what action should be taken. This may be done automatically or by the operator.

Supervision action and fault management – This is done depending on the danger of the diagnosed fault, which may be safe operation shut down, reliable operation which hinders further fault expansion. Reconfiguration which uses the standby components can also be done, Inspection, maintenance and repair to remove the fault.

In supervision and fault management, fault detection and diagnosis are very important for advanced methods for implementing such. The methods will be discussed in the following sections, which are essentially knowledge based fault detection and diagnosis.

2.4 Symptom Generation

Symptom generation is a way of processing data based on measured process variables that need to be performed to generate the first characteristic values [101]. The two ways of symptom generation are discussed in the following sections below.

2.4.1 Analytical Symptom Generation

The analytical knowledge about the process helps to produce quantifiable analytical information. These are as by R. Iserman (2005) [101];

Limit value checking of direct, measurable signals. The characteristic values are the violated signal tolerances.

Signal analysis of directly measurable signal models like correlation functions, frequency spectra, autoregressive moving average (ARMA) or the characteristic values like amplitudes, variances and frequencies or model parameters.

Process analysis by using mathematical process models together with parameter estimation, state estimation and parity equation methods. The characteristic values are parameters, state variables or residuals.

Sometimes special features can be extracted from characteristic values like physically defined process coefficients, or transformed residuals.

2.4.2 Heuristic Symptom Generation

This way of symptom generation uses qualitative information from human operators. This can be through human observation and inspection, heuristic characteristic values in the form of special noises, smells, color and the like are obtained.

The recordings of the process pertaining to maintenance, repairs, previous faults, life-time and load measures, make a big source of heuristic information. This is how heuristic symptoms are generated, which can be represented as linguistic variables as in small, medium or large or as vague numbers like, around this value [101].

2.5. Fault Diagnosis

Once the fault has been detected and identified, the next step is the fault diagnosis. The main task of fault diagnosis is in determining the size, type and the location where the fault has occurred [100].The time when the fault was detected can also be known.

Analytical and heuristic symptoms stated above are used in the fault diagnosis procedures. So it is important for these symptoms to be presented in the unified form. These can be like confidence numbers, membership functions of fuzzy sets or probability density functions after a statistical evaluation over some period of time. Then classification methods can be applied if the known pattern based procedure is preferred, to determine the faults from the symptom patterns. Otherwise more information of fault symptom relation, like in the form of logic fault symptom-trees or if then rules are known, reasoning methods with forward and backward chaining is applied.

CHAPTER 3

PRINCIPAL COMPONENT ANALYSIS

Principal component analysis (PCA) is a dimensionality reduction technique [9]. This is achieved by projecting the large data into a lower dimensional space that accurately characterizes the state of the process. This can greatly improve process monitoring procedures.

There are cases when data cannot be captured in two or three dimensions, in such cases methods have been developed to automate the process monitoring procedures. Application of PCA in such methods relies on three factors which are the production of lower dimensional representations by PCA of the data. This changes the data to become independent of the training data set, using the whole entire dimensionality of the observation space, hence improving the efficiency of detecting and diagnosing faults [10].

The structure made by PCA is useful in identifying either the variables that are responsible for the fault and/or the variables most affected by the fault as a second factor.

The third factor is that PCA can separate the observation space and into a subspace capturing the systematic trends of the process and a subspace containing essentially the random noise [9].

The general sensitivity of the process in monitoring scheme to faults can increase by applying one measure developed for one subspace. Another measure developed for the other subspace because it is widely accepted that certain faults primarily affect one of the two subspaces [9, 10]

It is in this chapter where PCA will be defined, discussion of different methods used for determining the order of the PCA presentation. Principal component analysis being used for fault detection, identification and diagnosis and a brief explanation of dynamic PCA (DPCA).

3.1 Definition of Principal Component Analysis

PCA as a dimensionality reduction technique is also an optimal method in capturing the variability of the data. Loading vectors are the orthogonal vectors determined by PCA which are ordered by the amount of variance as explained in the loading vector directions.

$$Max\ \boldsymbol{V^T X^T X V} / \boldsymbol{V^T V} \quad 3.1$$

Where $\boldsymbol{X}$ is a training data set of *n* observations and *m* process variables, so $\boldsymbol{X} \in R^{n \times m}$. The expression 3.1 above gives the loading vectors by solving the point of the optimization problem. The stationary points can be found by singular value decomposition and $\boldsymbol{V} \in R^{m}$ which gives;

$$\frac{1}{\sqrt{n-1}} X = U \sum V^T \quad 3.2$$

The above equation is equivalent to coming up with eigenvalue decomposition of the sample covariance matrix $\boldsymbol{S} \in R^{m \times m}$;

$$S = \frac{1}{n-1} \boldsymbol{X^T X} = V \Lambda V^T \quad 3.3$$

Where the diagonal matrix $\boldsymbol{\Lambda} = \sum^{\mathrm{T}} \sum \in \boldsymbol{R^{mxm}}$ and the eigenvalues in this diagonal matrix are in decreasing magnitude order ($\lambda_1 \geq \lambda_2 \geq \lambda_3 \ldots \geq \lambda_m$). This gives $\lambda_i = \delta_i^2$ which is the i^{th} eigenvalue which is equal to the square of the i^{th} singular value.

The crucial and most important part of the PCA is the selection of the loading vectors that are corresponding to the largest eigenvalues to be retained. This should be optimal in order to reduce the dimension of the original data but at the same time capture the variations of the data columns of the loading matrix $\boldsymbol{P} \in R^{mxa}$.This corresponds to the loading vectors associated with the first a singular values; the projections of the observations in $\boldsymbol{X}$ **(**original data which becomes $\boldsymbol{x} \in R^{nxa}$ after applying PCA procedure) into the lower dimensional space are contained in the score matrix,

$$\boldsymbol{T} = \boldsymbol{XP} \quad 3.4$$

Where $\boldsymbol{T} \in R^{n \times a}$ is projected into the m dimensional observation space.

$$\hat{\boldsymbol{x}} = \boldsymbol{TP^T} \quad 3.5$$

This gives the residual matrix $\breve{\boldsymbol{x}} = \boldsymbol{x} - \hat{\boldsymbol{x}}$ 3.6

where $\hat{\boldsymbol{x}}$ and $\breve{\boldsymbol{x}}$ are the same dimensions as $\boldsymbol{x}$**.**

It is in the residual matrix where the variations in the observation space are put together in the loading vectors associated with the $(m - a)$ smallest singular values. Vector $\hat{\boldsymbol{x}}$ spanned subspace is the score space; it is this space that has the process information [9].

A new observation vector in the testing set $\boldsymbol{x} \in R^m$ can be projected into the lower dimensional score space $\boldsymbol{t_i} = \boldsymbol{x^t p_i}$, where $\boldsymbol{p}_i$ is the i^{th} loading vector and $\boldsymbol{t_i}$ is a vector in **T** matrix. The transformed variable $\boldsymbol{t}_i$ is called the i^{th} principal component of $\boldsymbol{x}$ [1, 9].

The difference between the transformed variables and the transformed observations is that, the transformed variables are called the principal components while the transformed observations are the scores [38]. Figure 3.1 below shows the PCA from multidimensional data, this is the key component for fault detection [9].

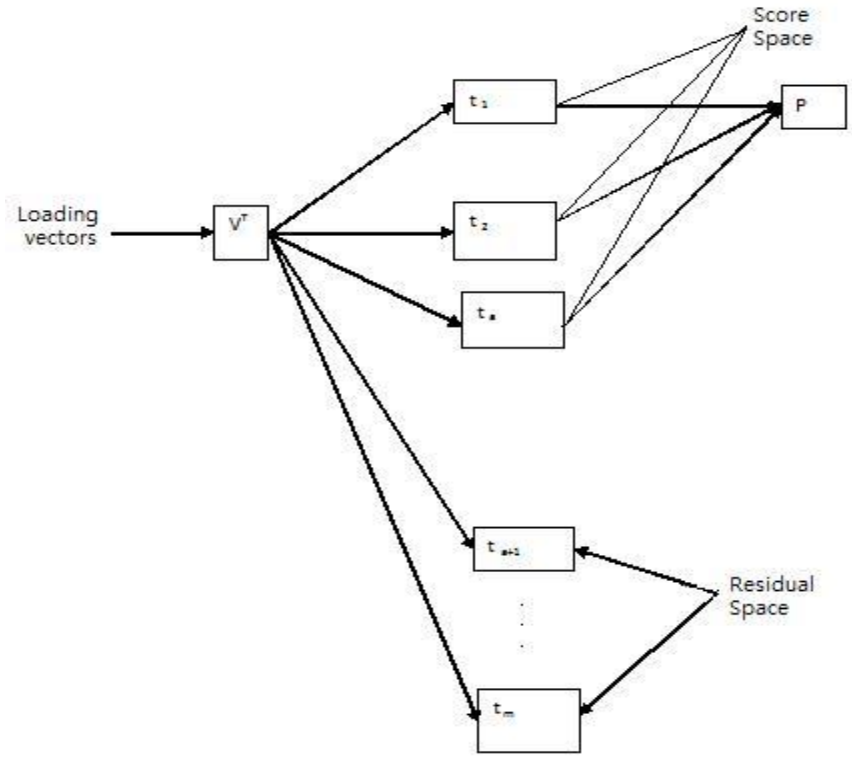

Figure 3.1: The projection of the observation vector $\boldsymbol{x}$ into the score and residual spaces and filtered data $\hat{\boldsymbol{x}}$.

The total variance of $\boldsymbol{x}$ projected along $\boldsymbol{V}$ is equal to the trace of $\boldsymbol{\Lambda}$, the i^{th} value in the diagonal $\boldsymbol{\Lambda}$ indicates the amount of variance for that particular i^{th} principal component. The score matrix $\boldsymbol{T}$ is then calculated from equation 3.4.

3.2 Reduction Order

Determination of the number of principal components to be retained when using PCA as a method of dimensionality reduction is the most important. It is the common fact that PCA space corresponding to

the larger eigenvalues or singular values describe the most information or the systematic or state variations occurring in the system. The PCA space with the less significant eigenvalues or singular values describes the random noise [38].

The number of principal components to be retained for PCA will help to analyze that part which carries the systematic or state variation and that part with random noise separately, and this is called reduction order.

There are a number of techniques that are used to determine the value of the reduction order as described in [12, 39, and 40].The following are some of the most commonly used methods.

3.2.1 Cumulative Percentage Variance

Cumulative percentage variance method is one of the commonest methods used in determining the optimal number of PCs to be retained when using PCA for fault detection.

This involves selecting the cumulative percentage of the total variation which is desired for the selected PCs should contribute, for example one can choose 85% or 90% of the total variation. The required number of the PCs is then the smallest value for which this chosen percentage exceeded.

For example if a number of PCs are considered then

$$CPV(a) = 100\left[\frac{\sum_j^a \lambda_j}{\sum_j^m \lambda_j}\right]\% \quad 3.7$$

Where m the total number of eigenvalues and with this, one is selects the desired CPV such as 85% or 90%, which is subjective. This makes the CPV method to be somehow ambiguous because the CPV is monotonically increasing with the number of PCs retained. Since this minimum percentage is chosen arbitrarily, it may be too low or too high for that particular application [78].

3.2.2 Scree Test Method

The simplicity of this method makes it easy to be used by most researchers. It was discussed and named by Cattell [84].

This method involves plotting the eigenvalues against the number of principal components; these eigenvalues are arranged in ascending order and linked with a line (as an example see figure 3.1). The number of PCs to be retained are decided at which the value is found on the slope of the lines joining the plotted points ‘steep’ to the left of the PCs and not on the ‘steep’ right of the same (in case of the example Scree plot the number of PCs to be retained is 3).

Example of the Scree Plot

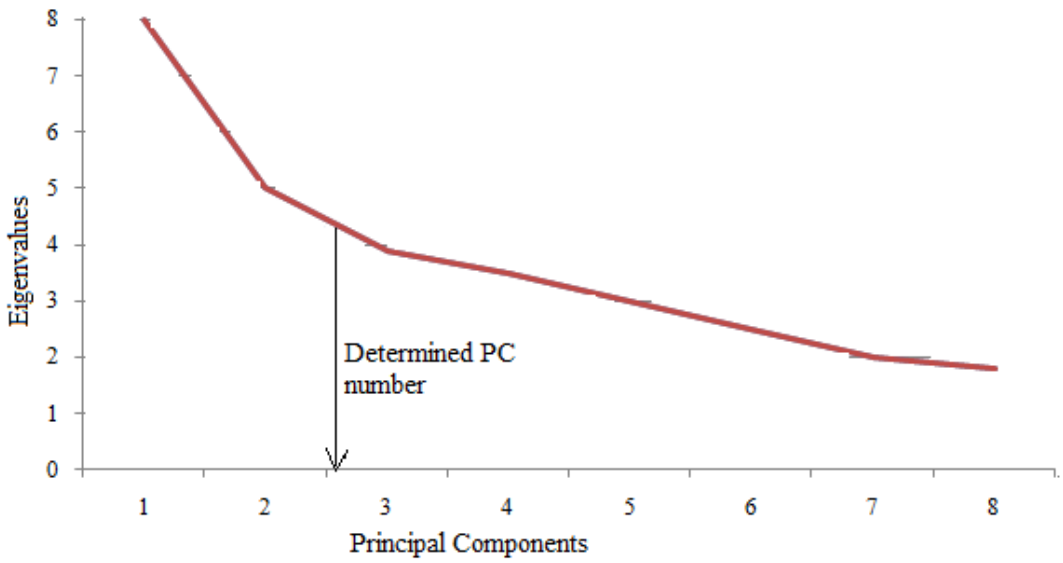

Figure 3.2: The Scree Plot for PC Determination

The graph is examined to determine the point at which the least significant drop or break takes place, in other words, where the line almost levels off. The logic behind this method is that, this point divides the major components with important information from the minor components with noise [38].

Most of the times the eigenvalues are small in numerical values, some scientist prefer plotting the logarithm values of the eigenvalues against the number of principal components. Then the same way of determining the optimal number of PCs to be retained is applied [39].

Since this comes as a curve, it is ambiguous to determine the exact number; this makes it difficult to automate this method of reduction order [9].

Sometimes more breaks from linearity occur in the profile and it is in situations like these that the Scree test method becomes difficult to use.

3.2.3 Parallel Analysis

Horn developed Parallel Analysis (PA) method as a modification of Cattells' Scree diagram to alleviate the component indeterminacy problem.

Basically the PA method is built on PCA models of two matrices; one is the original data matrix (line plot 1 in figure 3.3) and the other is an uncorrelated data matrix (line plot 2 in figure 3.3) with the same size as

the original matrix. When the eigenvalues of each matrix are plotted in the same figure, all the values above the intersection point represent the process information and the values under this point are considered noise [21]. Since we have the exact point to get the number of PCs to be retained (in the example PA plot 3 is the number of PCs to be retained), the PA method is not ambiguous. This method can be automated and is the most recommended method because of its best overall performance.

The Parallel Analysis Example Plot

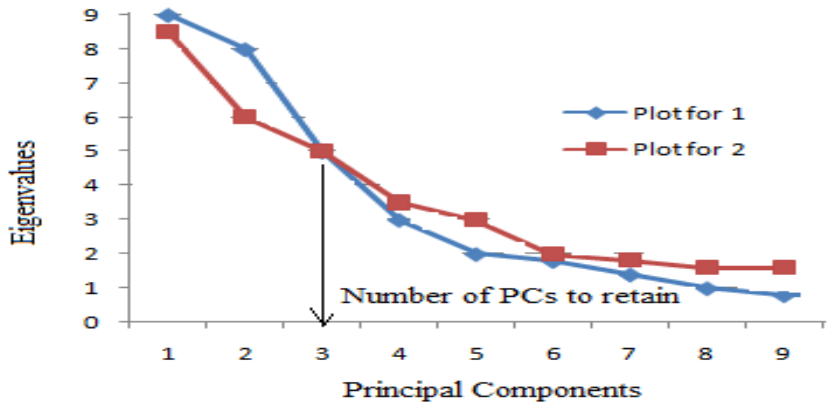

Figure 3.3: The Parallel Analysis Plot for PC Determination

3.2.4 Fault Signal to Noise Ratio (SNR)

This is one of the most recent methods, introduced in 2009 [22], of determining the reduction order for fault detection in PCA. The method is based on the proposed fault signal to noise ratio (fault SNR). It indicates the relationship between the sensitivity of fault detection and the number of PCs.

Where the fault SNR is put to the maximum, the optimal number of PCs can be selected and the performance of fault detection improved [92]

This allows the retention of any number of PCs on its own and observes how sensitive it is to fault detection.

It is the interest of the author of this dissertation to compare the performance of this modern method to the other three common old methods of determining the reduction order and draw conclusion on the strengths and the weaknesses of this method compared to the others, which is given in chapter 6.

3.3 Fault Detection

3.3.1 T^2 – Statistic

As expressed in the above section, PCA can also be used as a technique for fault detection. The calculation of the T^2 statistic can help achieve this. For observation vector $\boldsymbol{x}$ and $\boldsymbol{\Lambda} = \boldsymbol{\Sigma^T \Sigma}$ is invertible, T^2 statistic can be calculated as;

$$T^2 = \boldsymbol{x^T v(\Sigma^T \Sigma)^{-1} v^T x} \quad 3.8$$

Where $\boldsymbol{V}$ is taken from the covariance matrix $\boldsymbol{S}$, as in $\boldsymbol{S = V^T V}$ and $\delta_i^2 = \boldsymbol{\Sigma^T \Sigma}$ which are equal to the diagonal elements of the matrix $\boldsymbol{\Lambda}$.

When the number of observation variables is large and the amount of data available is relatively small, this is like the number of rows is large and the number of columns is relatively small, the T^2 statistic tends to be an inaccurate presentation of the in-control process behavior. This is more in the loading vector directions corresponding to the smaller singular values. These inaccuracies may have a big effect on the calculation of the T^2 statistics; this is so because, the squares are inverted as indicated in the equation 3.8 above.

As a way of avoiding this problem, the loading vectors associated with the largest singular values should be used for the calculation of the T^2 statistic.

Including the $\boldsymbol{P}$ matrix, the loading vectors with a number of largest singular values; the T^2 statistic for the lower dimensional space becomes [42];

$$T^2 = \boldsymbol{x^T P} \sum_a^{-2} \boldsymbol{P^T x} \quad 3.9$$

Where $\boldsymbol{\Sigma_a}$ has the first a rows and columns of the $\boldsymbol{\Sigma}$. As it is T^2 only measures the variations in the score space. If the actual mean and covariance are known the;

$$T_\alpha^2 = X_\alpha^2(a) \quad 3.10$$

With the known actual covariance matrix from the sample covariance matrix, the T^2 statistic threshold becomes;

$$T_\alpha^2 = \frac{(a(n-1)(n+1))}{(n(n-a))} F_\alpha(a, n-a) \quad 3.11$$

Where F_α can be calculated from the table in appendix C. To come up with the outliers in the training set, the threshold becomes;

$$T_\alpha^2 = \frac{(n-1)(n-1)(a/(n-a-1))F_\alpha(a,n-a-1))}{n(a/(n-a-1))F_\alpha(a,n-a-1))} \quad 3.12$$

The T^2 in equation (3.7) is not affected the inaccuracies in the smaller values of the singular values, hence better represents the normal behavior of the process. Now this can be used for fault detection when compared with T^2 from equation (3.6). This is because the T^2 statistic is the measurement of the systematic variations of the process, and any violation of the threshold would indicate that the fault occurred because the systematic variations are not within the threshold.

3.3.2 Q – Statistic

The portion of the residual space which corresponds to the smaller singular values can be monitored efficiently by the use of Q – statistic [5, 39, and 41].

$$Q = \boldsymbol{r}^T\boldsymbol{r} \quad 3.13$$

$$\boldsymbol{r} = (\boldsymbol{I} - \boldsymbol{P}\boldsymbol{P}^T)\boldsymbol{x} \quad 3.14$$

The residual vector $\boldsymbol{r}$ is the projection of the observation $\boldsymbol{X}$ into the residual space. The Q – statistic measures the total sum of variations in the residual space. The Q – statistic is not over sensitive to the inaccuracies in the smaller singular values [5]. The Q – statistic is sometimes known as the Squared Prediction Error (SPE), which is the squared 2 – norm measuring the deviations of the observations to the lower dimensional PCA presentation.

The distribution of the Q – statistic can be approximated as;

$$Q_\alpha = \theta_1\left[\frac{h_0 c_\alpha \sqrt{2\theta_2}}{\theta_1} + 1 + \frac{\theta_2 h_0(h_0-1)}{\theta_1^2}\right]^{\frac{1}{h_0}} \quad 3.15$$

Where $\theta_i = \sum_{j=a+1}^{n} \delta_i^2$ 3.16

And $h_0 = 1 - \frac{2\theta_1\theta_3}{3\theta_2^2}$ 3.17

and C_α is the Q – statistic measures the random variations of the process while the threshold Q_α is applied to define the normal variations of the random noise, and any violation of the threshold can indicate that the random noise has significantly changed.

The Q – statistic and the T^2 statistic with their thresholds would indicate or detect different types of faults. The advantages of both can be used if both of the techniques are together.

3.4 Fault Identification

Fault identification is a step that follows fault detection. After realizing that a fault occurred in the system, there is need to identify the fault and where it happened, knowing the actual variable that is affected. Univariate statistical techniques are usually used for fault identification, the observation vector **x** with normalized errors for each variable x_j;

The error $e_j = \frac{(x_j - u_j)}{s_j}$ 3.18

where u_j is the mean and S_j is the standard deviation of the j^{th} variable.

There a disadvantage of using the univariate statistical techniques for fault identification, that is, it can leave out variables that are responsible for the fault. This is so because the techniques do not account for correlations among the process variable [6]. To deal with this problem PCA, contribution plots are a PCA approach to fault identification, these plots takes into account the special correlations and this improves the effectiveness of the univariate statistical techniques for fault identification [6, 7].

There is a procedure that is followed to a T^2 violation;

Check the normalized scores (t_i/δ_i^2) for the observation $\boldsymbol{x}$ and determine the $r \leq a$ scores responsible forout of control status.

Calculation of each variable (x_j) to the out of control variable t_i

$$cont_{i,j} = (ti/\delta_i^2)p_{i,j}(x_j - u_j);$$

Where $p_{i,j}$ is the $(i,j)th$ element of loading matrix P.

When $cont_{i,j}$ is negative, set it to equal to zero.

Calculate the total contribution of the jth process variable, x_j.

$$Cont_j = \sum_{i=1}^{r}(cont_{i,j}).$$

Plot $Cont_j$ for all m process variables, x_j, on a single graph.

The total contribution $Cont_j$ can help to give a priority to the variables for the fault. So those with high $cont_j$ value would be given the first priority followed by those with least $cont_j$values.

Wise et al [19] came up with a PCA approach to fault identification which quantifies the total variation of each of the process variables in the residual space.

$$\hat{S}_j^2 = \sum_{i=a+1}^{p} p_{i,j}\, \delta_j^2 \qquad 3.19$$

$$\frac{s_j^2}{\hat{s}_j^2} > F_\alpha\,(q - a - 1, n - a - 1) \qquad 3.20$$

$F_\alpha\,(q - a - 1, n - a - 1)$ Is the (1 - α) percentile limit using the F distribution [37] (appendix C).

If $s_j^2 \neq \hat{s}_j^2$

$$\frac{s_j^2}{\hat{s}_j^2} > \frac{F_\alpha}{2}\,(q - a - 1, n - a - 1) \qquad 3.21$$

Or $$\frac{\hat{s}_j^2}{s_j^2} > \frac{F_\alpha}{2}\,(n - a - 1, q - a - 1) \qquad 3.22$$

But a large shift in the mean inside the residual space occurs if [8, 37];

$$\frac{u_j - \hat{u}_j}{\hat{s}_j\sqrt{(\frac{1}{q-a}+\frac{1}{n-a})}} \geq \frac{t_\alpha}{2}(q + n - 2a - 2) \qquad 3.23$$

Or

$$\frac{u_j - \hat{u}_j}{\hat{s}_j\sqrt{(\frac{1}{q-a}+\frac{1}{n-a})}} < \frac{-t_\alpha}{2}(q + n - 2a - 2) \qquad 3.24$$

u_j And $\hat{u}_j$ are the means for x_j for new and training set respectively and $\frac{t_\alpha}{2}(q + n - 2a - 2)$ is the (1 - α/2) percentile limit using the t distribution.

Detecting the variables responsible for the out of control status using the Q – statistic, can be achieved by equations (3.18), (3.22) and (3.23) and to add on that the variables can also be prioritized by using the expression values of the same equations [9, 10, 11].

A fault identification measure that is based on an observation vector at a single time instant is the normalized error;

$$RES_j = \frac{r_j}{\hat{s}_j} \quad 3.25$$

In the above equation r_j is the j^{th} variable of the residual vector. This can also be of use in the prioritizing of the variables, such that that with the highest normalized errors is given priority.

3.5 Fault Diagnosis

For a fault to be properly diagnosed it may take sometime for engineers and operators to do it using the fault identification methods.

Employing automated fault diagnosis scheme is one of the ways to reduce the time engineers may take to diagnose a fault that has occurred in the process.

One of the ways is to construct separate PCA models for each process unit [12]. This can help a big deal because a fault associated with a particular process is assumed to occur if the PCA model for that unit indicates that the process is out of control. It distinguishes the fault isolation techniques which are based on non – supervised classification from fault diagnosis techniques which are based on supervised classification.

Amongst the many techniques that use PCA for fault diagnosis, the simplest is the construction of a single PCA model that defines regions of in the lower dimensional space. This classifies whether a particular fault has occurred [8], the only drawback is that, this approach is no effective when a significant number of faults has occurred [13].

The other PCA approach which can handle a larger number of faults than using a single PCA model is to come up with a PCA model. This model which is based on data collected during each specific fault situation. The application of Q and T^2 statistics [13, 14, 15, and 16] to each PCA model can predict which faults that have occurred. This approach is a combination of PCA and discriminant analysis [15].

Developing a model based on the data from all fault classes is another way of using PCA for fault diagnosis.

$$g_i(x) = -1/2(x - \bar{x}_i)P\,(P^T S_i P)^{-1}\,P^T(x - \bar{x}_i) + ln\,(p_i) - 1/2ln[det(P^T S_i P)] \quad 3.26$$

Where $\bar{x}_i$ is the mean vector;

$$\bar{x}_i = 1/n_i \sum_{x_j \epsilon X_i} x_j \quad 3.27$$

n_i Is the number of data points is fault class 1, X_i is the set of vectors X_j which belongs to the fault class i and $S_i \epsilon R^{mxm}$ in the sample covariance matrix.

The discriminant analysis techniques that are used with the multiple PCA models are the score discriminant, residual discriminant and combined discriminant [14]. If the PCA models only retain the important variations in discriminating between the faults, then equation (3.27) becomes

$$g_i(x) = -1/2x^T\ P_{\ i} \sum_{a,i}^{-2} P_i^T x - 1/2ln[det(\sum_{a,i}^{2})] + ln(p_i) \quad 3.28$$

Where p_i is the overall likelihood of fault class I, [16 and 43].When the important variations in discriminating between the faults are contained in the residual space for each fault class. The observation is represented by the fault class i with the minimum residual discriminant.

$$\frac{Q_i}{(Q_\alpha)_i} \quad 3.29$$

But the important variations is discriminating between the faults are combined both within the score and residual space, then an observation is likely to be represented by the fault class I with the minimum combined discriminant [9].

$C_i[\frac{T_i^2}{(T_\alpha^2)_i}] + (1 - C_i)[\frac{Q_i}{(Q_\alpha)_i}]$ where C_i is the weighing factor between 0 and 1.

All the equations (3.23, 3.24 and 3.26) can be used to diagnose a fault assuming an out of control observation does not represent a new fault.

It is important to assess the likelihood of successful diagnosis. Quantitative measure of similarity between the covariance structures of two classes [15, 16], this measure is called similarity index, for the two classes calculated as;

$$F = \frac{1}{m}\sum_{j=1}^{m} \tilde{\sigma}_j \quad 3.30$$

Where $\tilde{\sigma}_j$ is the j^{th} singular value of $V_1^T V_2$ and matrices V_1 and V_2 contain all m loading vectors for both classes. The value of F ranges from 0 to 1, where 0 means lack of similarity and 1 indicates exact similarity [17]. The similarity index can be applied to PCA models by replacing V_1 and V_2, with loading matrix $\mathbf{P_1}$ for class 1 and loading matrix $\mathbf{P}_2$ for class 2 respectively.

It is important to have a method for analyzing a control system that takes into account the correlations in the data.

This is where dynamic principal component analysis comes in.

To make sure that correlations are present in the data, autocorrelation chart of the principal components is used [16, 19].

EWMA/CUSUM charts with PCA are one of the methods for establishing the correlation in the data beside the approach of averaging the measurement over a number of data points. PCA can also consider the serial correlations by augmenting each observation vector with the previous *h* number of observations and come up with a matrix that looks like the one below;

$$X(h)=\begin{bmatrix} x_t^T & x_{t-1}^T & \cdots & x_{t-1}^T \\ \vdots & \vdots & \ddots & \vdots \\ x_{t+h}^T & x_{t+h-n-1}^T & \cdots & x_{t-n}^T \end{bmatrix} \quad 3.31$$

Where X_t^T is the m – dimensional observation vector in the training set at time interval t. For single input single output processes (SISO), it is described as ARX (h) model, where

$$Y_t=a_1 y_{t-1}+\ldots\alpha_h y_{t-h}+\beta_0 u_t+\beta_1 u_{t-1}+\ldots\beta_h u_{t-h}+e_t. \quad 3.32$$

Here e_t is the white noise and α and β are constants coefficients.

$$X_t^T=[Y_t \quad U_t] \quad 3.33$$

And this makes equation (3.31) to become;

$$X(h)=\begin{bmatrix} y_t & u_t & y_{t-1} & u_{t-1} & \cdots & y_{t-h} & u_{t-h} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ y_{t+h-n} & u_{t+h-n} & y_{t+h-n-1} & u_{t+h-n-1} & \cdots & y_{t-n} & u_{t-n} \end{bmatrix} \quad 3.34$$

The ARX (h) model shows that the first x (h) columns to be linearly related to the remaining columns [9]. When PCA is applied on x (h) the eigenvector corresponding to the zero eigenvalue would reveal the ARX (h) correlation structure [21].

The eigenvector corresponding to nearly zero eigenvalue will be an approximation of the ARX (h) correlation structure [9, 21, and 22].

So the application of PCA in the matrix shown in (3.29) is what is called the Dynamic principal component analysis (DPCA).The application of fault detection techniques and fault diagnosis for static PCA can also be applied on dynamic PCA.

CHAPTER 4

THE LABORATORY THERMAL SYSTEM

The thermal system that has been in the control lab for many years is used for various experiments especially for the design of the controllers by researchers. This shows how important the device is to Control Laboratory and hence the need for it to be in good working condition. Figure 2.1 below has the picture of this laboratory thermal system which is used validation of the techniques discussed in this dissertation.

Figure 4.1: UCT laboratory thermal system

There are various laboratory thermal systems of the same nature that could have been used for this purpose. The only main difference is that some have a single input and single output (as the one in figure 4.2) while the one used has two inputs and two outputs which can be used for MIMO experiments. That

means will have four variables and with the application of PCA could be reduced to one, two or three variables while retaining the original signal information.

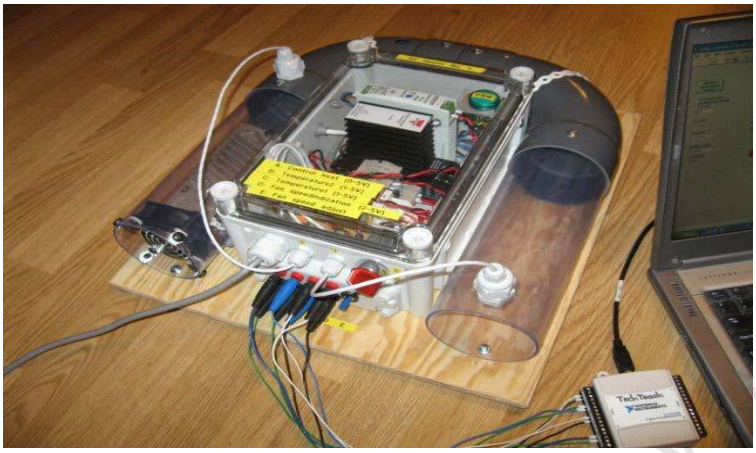

Figure 4.2: SISO Laboratory Thermal System

4.1 System Definition

The thermal system controls the heating and cooling by adjusting heat transfer based on the temperature measurement and the desired 'set-point' temperature. It comprises of two heat blowers and two heat sensors.

The blower's heat goes up or down depending on the input voltage. The heat is the input to the system, while the temperature of the air at the thermal sensors represents the output of the system. These sensors are the transducers that provide an electrical response that is proportional to the temperature.

The two sensors are positioned in such a way that the heat from the heater 1 directly affects sensor 1 and partly affects sensor 2. The heat from heater 2 directly affects sensor two and partly affects sensor 1.

In any control problem it is important to know the characteristics of each component so that the behavior is known for different situations. For example if a voltage is applied to the electrical heater, a certain amount of heat is provided over time usually reaching a steady state heat output for a fixed applied voltage.

For the thermal system that is used in this thesis, the concept stated above applies. Step input voltage of 1V was applied for the input voltage. The minimum voltage was 4.5V and the maximum was 6.5V. The output voltage was related to the input voltage in that when the input voltage increases the output voltage decreases. The electric coils heats up with a decrease in the input voltage and this increases the output voltage from the temperature sensors. Also the coils cool down with an increase in the input voltage that

decreases the output voltage from the sensors. The lower constraint on the input voltage of 4.5V was imposed for the safety of the heater coils, protecting them from too much heat that may damage the heater casing.

As there are two heaters and two sensors the system is multivariable (MIMO) and results in the 2 x 2 matrix of the form given below;

$$G_{Full} = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} \quad 4.1$$

In the structure above, g_{11} is the transfer function for relating input 1 to output 1, while g_{12} is for input 1 and output 2. The same definition is true for g_{21} and g_{22}.

4.2 The Identification of Thermal System Model

The identification of the system model is necessary before controller design and application of fault detection technique. There are a number of system identification methods [Appendix B]. One of identification of the system model method Liu and Gao 2010 [107], they proposed frequency response domain step response identification to obtain a system model of a continuous-time process with time delay. In this method the damping factor is introduced to create a Laplace transform which enables the frequency response estimation algorithm to be proposed. The proposed algorithm can find the optimal fitting accuracy over the frequency response range that is of interest for the control design. It was concluded that this method shows good robustness with respect to various choices of the damping factor that is used in the computation. This method is also suitable for a system with noise and is applicable to higher order systems.

The other method proposed by Lee and Edgar 2010 [108], is a graphical method that uses an integral of the impulse response to identify the first order model and the time delay of the system model. The advantage of this method is that it does not disturb the process and the integral of the response that makes the method applicable to noisy pulse and noise step responses with time delay.

In this dissertation used step test is used. The linear model of the dynamic behavior of the thermal system is obtained using the step test because of its simplicity and wide application.

4.2.1 Thermal System Step Test

The step test data were obtained relating to the two inputs and two outputs of the laboratory thermal system for an open loop system. The figure 4.1 has the step tests for the thermal system.

The thermal system is modeled using a first order model in which the relationship between the output and the input voltages are indicated in the equations below. The change in the system output voltage Δy and change in the system input voltage Δu gives the system gain A. Time constant T of the process which is 63% of the system output settling time.

$$A = \frac{\Delta y}{\Delta u} \qquad 4.2$$

Four models were calculated and the following equations gives g_{11} calculated from output 1 and input 1 voltage, g_{22} gives the model from output 2 and input 2. g_{12} , g_{21} are the models of the interaction of the system inputs and outputs. The system data that was used to come up with the models are given in the following figures;

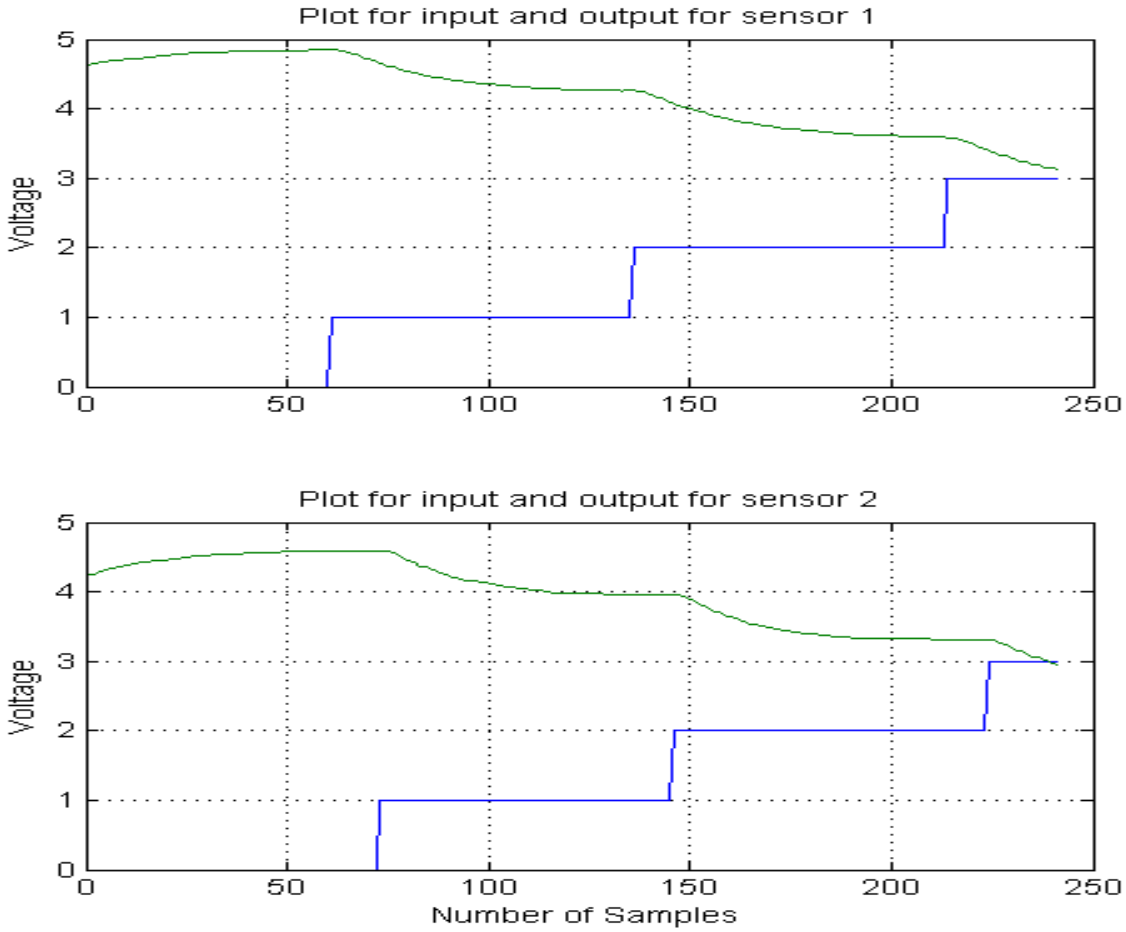

Figure 4.3: The unit step response for the two thermal sensors.

The models from this data were found to be:

$g_{11} = \frac{0.067}{s+0.132}$

$g_{22} = \frac{0.057}{s+0.132}$

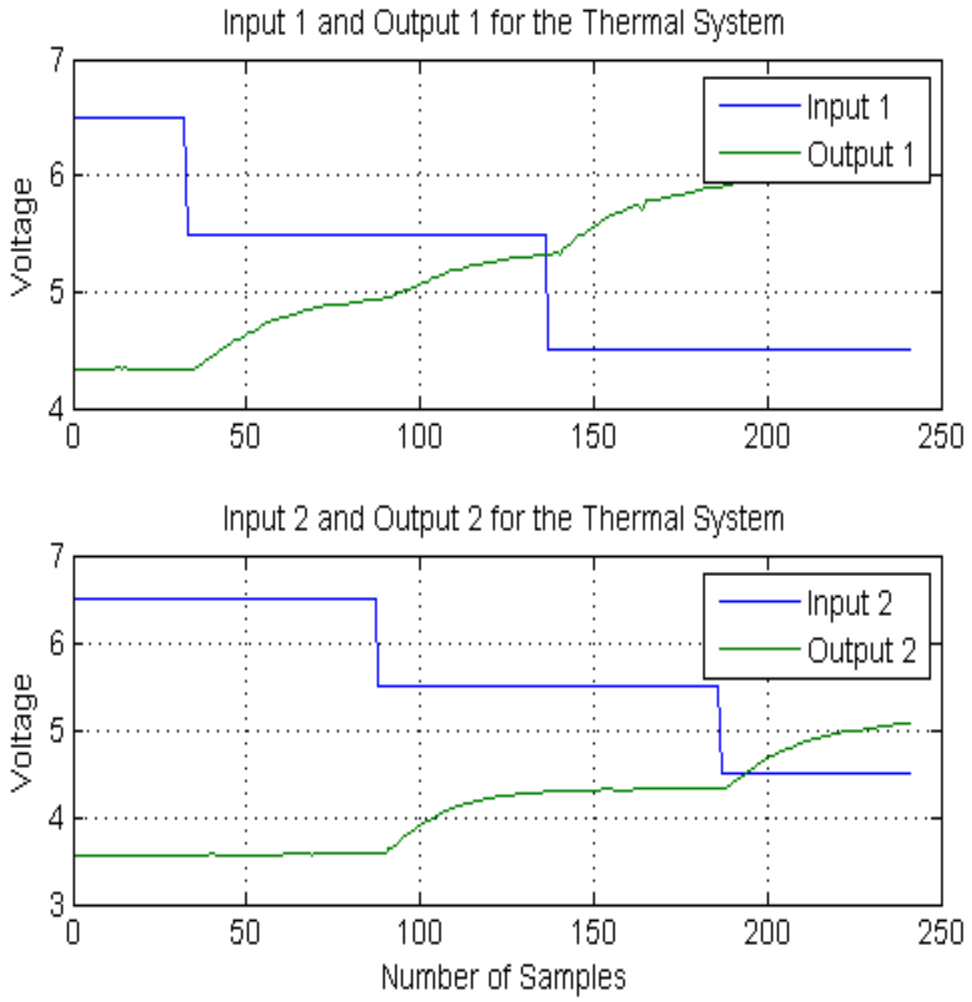

Figure 4.4: The unit step response showing the interaction of the two thermal sensors

The models from this data were found to be:

$g_{12} = \frac{0.00053}{s+0.132}$

$g_{21} = \frac{0.0035}{s+0.132}$ 4.3

As seen on the equations above, the two most significant models are g_{11} and g_{22} hence they are the ones to be discussed in detail. The validation and stability tests will be applied on these two.

4.2.2 Step Perturbation

In the experiments that were conducted for the normal performance of the thermal system, the step tests were used to come up with the thermal system models. The figure below has the step response of the thermal system.

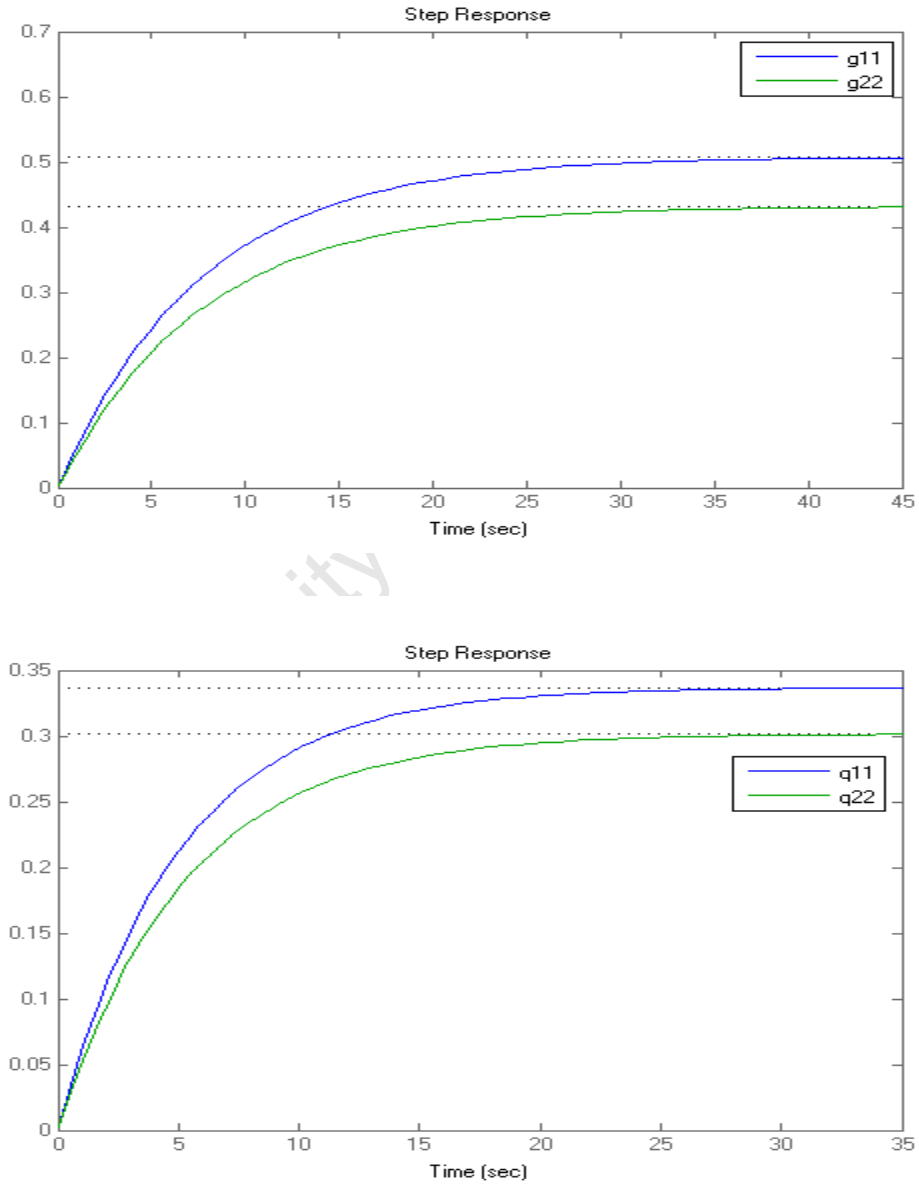

Figure 4.5: The step response for the thermal system at a sampling rate of 0.4 seconds.

However the sampling interval used was 0.4 second; this was chosen to easily note any changes taking place in the system when used for fault detection.

4.2.3 Validation of the System model

In order to implement the control algorithms and the fault detection, a matched model is needed. Any mismatch between the real system and the model may lead to a failure to meet the design specification that may cause instability in the closed loop system. Hence the importance to ensure that the model is valid, by using the step built in model using Matlab function from math-works, two dataset were generated for comparison. The real time output data and model generated data were compared against each other. The data comparison results are shown in the figures below.

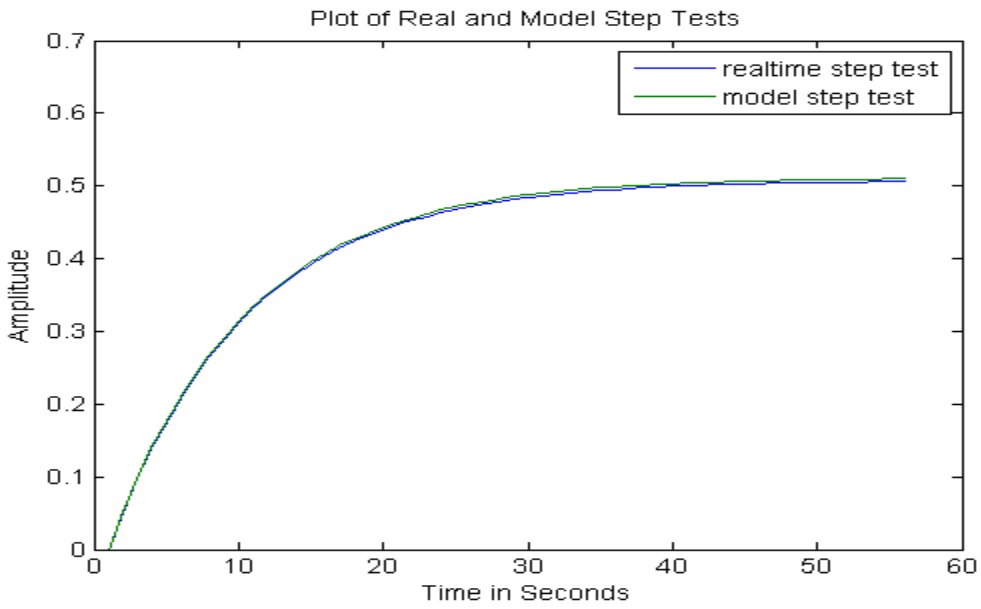

Figure 4.6: Validation step response plot for g_{11} system models

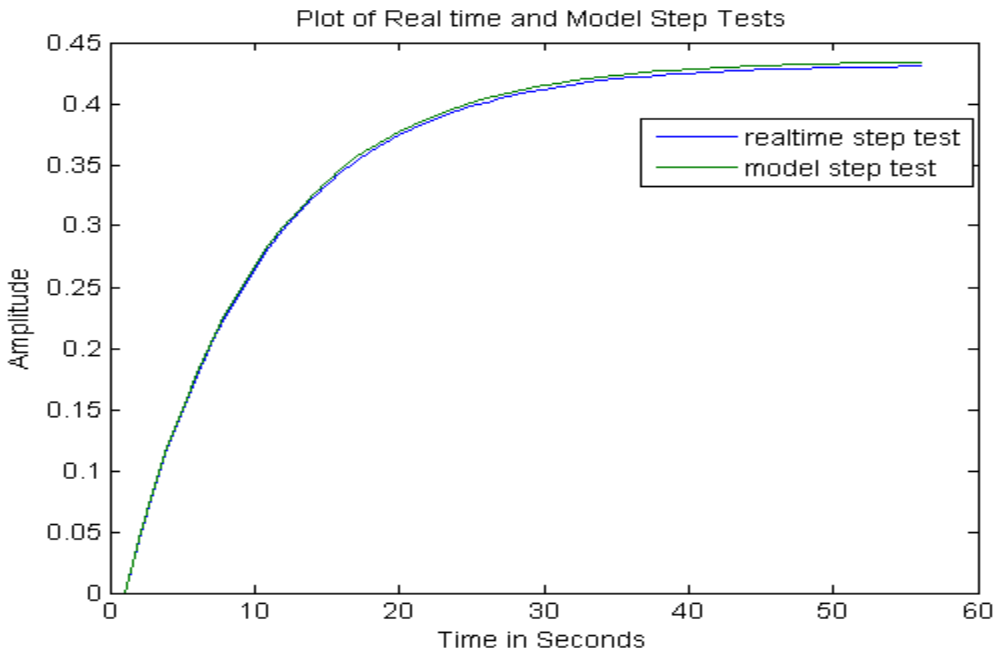

Figure 4.7: Validation step response plot for g_{22} system models.

The validation of step response for the real time and the model of the system results show that system can be used for the experiments needed in this dissertation.

The system real data was used for the experiments for fault detection in this dissertation.

4.3 MIMO Structure Analysis

4.3.1 General Analysis

In order to achieve the required results of the laboratory thermal control system, it should be possible to measure outputs that can give the true reflection of the amount of heat in the heater coils through the sensors.

The two sensors are placed in a position not too close to its respective heater blower, because this may make the sensor not to get the effect of the other heater blower. In this selected position, each sensor gets the effect of the two blowers, but gets more of its respective blower.

This can be shown to result in two fully coupled MIMO system with a structure of a form as below;

$$G(s) = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} \quad 4.4$$

In the transfer function matrix above, a transfer function g_{ij} (on the i^{th} row and j^{th} column) is from the i^{th} input and the j^{th} output. The geometrical analysis of this system shows that the DC gains of $\{g_{11}$ and $g_{22}\} > \{g_{12}$ and$g_{21}\}$, and depends on the distance between the heater blowers and the sensors.

$$G(s) = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} = \begin{bmatrix} \frac{0.067}{s+0.132} & \frac{0.00053}{s+0.132} \\ \frac{0.0035}{s+0.132} & \frac{0.057}{s+0.132} \end{bmatrix} \quad 4.5$$

4.4 Controller designing

Putting the two loops in feedback with a unit feedback element and a unit controller the closed loop models become; The closed loop models for the most dominant models g_{11} and g_{22} with a unit feedback becomes;

$$q_{11} = \frac{0.067s + 0.008844}{s^2 + 0.331s + 0.02627} \quad 4.6$$

$$q_{22} = \frac{0.057s + 0.007524}{s^2 + 0.321s + 0.02495}$$

The thermal controller models were used for designing the controller using the root locus method. The most significant models were used for this, G_{11} and G_{22}. The controller K_1 and K_2 are for G_{11} and G_{22} respectively.

4.4.1 Root Locus Method

This control design method analyses dynamic systems based on the basis of their pole and zero positions in the s-plane. It gives the considerable guidance as to where the poles and zeros of the compensating element K(s) should be placed for improving the dynamic behavior of the overall system [96]. In a nutshell, the positions of system poles and zeros in the s-plane give a good indication of how that system will perform in the time domain.

Using the gain of a compensator, and practicing trial and error, the author came up with the following controller K_1 and K_2 which are the same because the system models are almost the same.

$K_1 = k\frac{s+2.5}{s}$

$G_{11}(s) = K_1 \times \frac{0.067}{s+0.132}$ 4.7

$K_2 = k\frac{s+2.5}{s}$

$$G_{22}(s) = K_2 \times \frac{0.057}{s+0.132} \quad 4.8$$

Where k = 255 is the controller gain for G_{11} and G_{22} as seen on the root locus plot. This is slightly different at the same frequency of approximately 3rad/sec. The system is stable as the poles are on the left hand side of the s- plane and that's how K_1 and K_2 was chosen.

Following are different values of K_1 and K_2 that were considered to get to the one used.

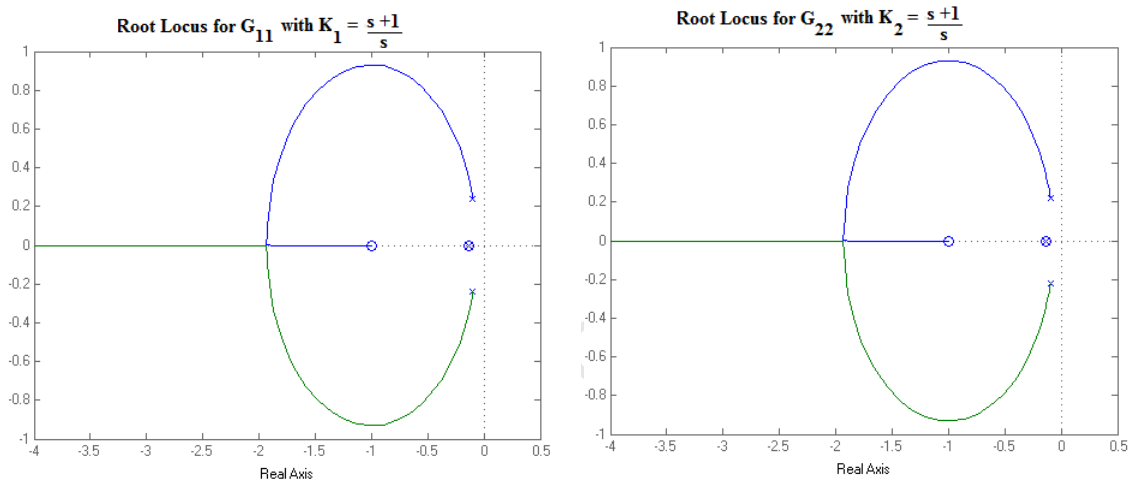

Figure 4.8: Root locus plots for G_{11} and G_{22} with K = $\frac{s+1}{s}$

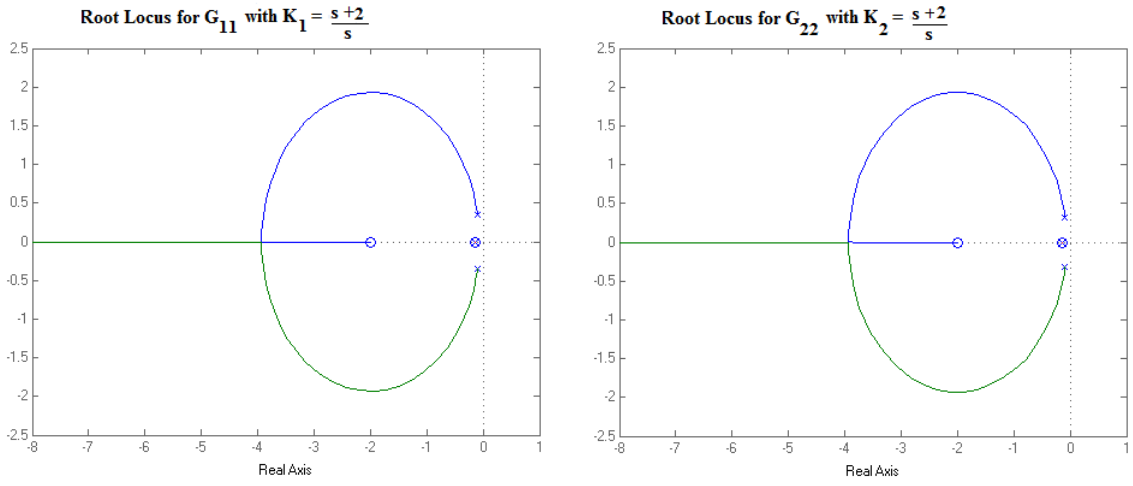

Figure 4.9: Root locus plots for G_{11} and G_{22} with K = $\frac{s+2}{s}$

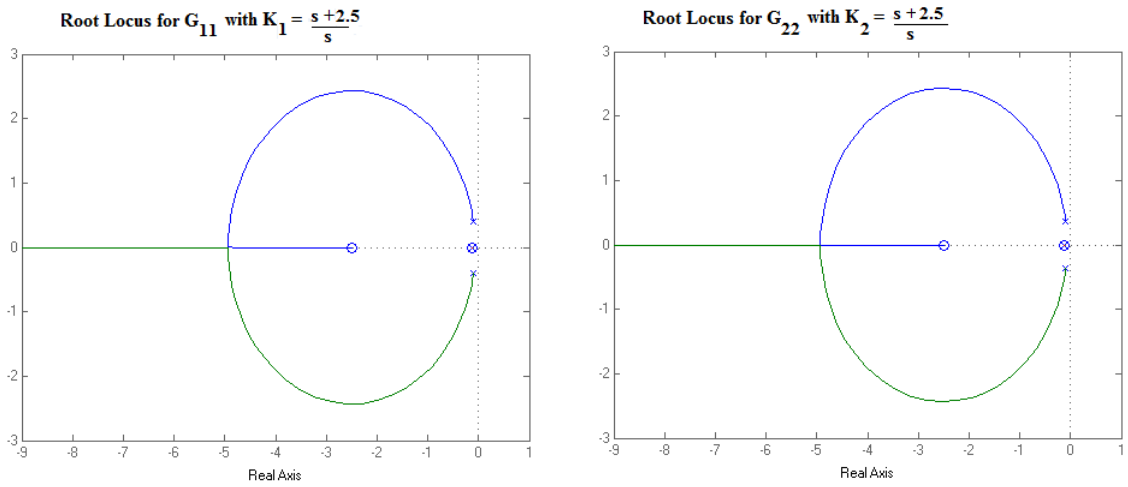

Figure 4.10: Root locus plots for G_{11} and G_{22} with K = $\frac{s+2.5}{s}$

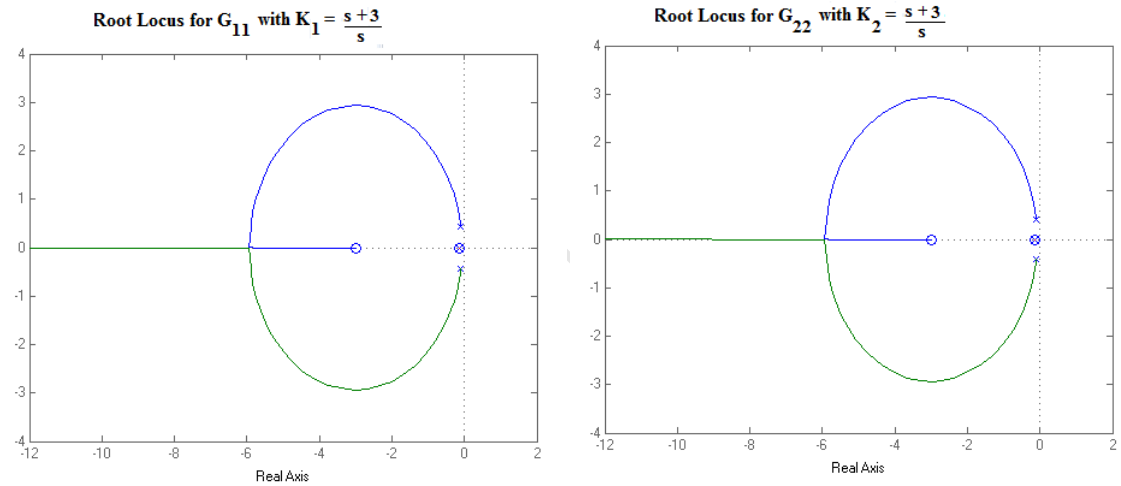

Figure 4.11: Root locus plots for G_{11} and G_{22} with K = $\frac{s+3}{s}$

When the physical model of the thermal system is positioned in a way that the sensor 1 is moved close to the heater blower 1 to eliminate g_{12} the two thermal control systems can be designed. They can be cascaded with two feedback loops as shown in the diagram below.

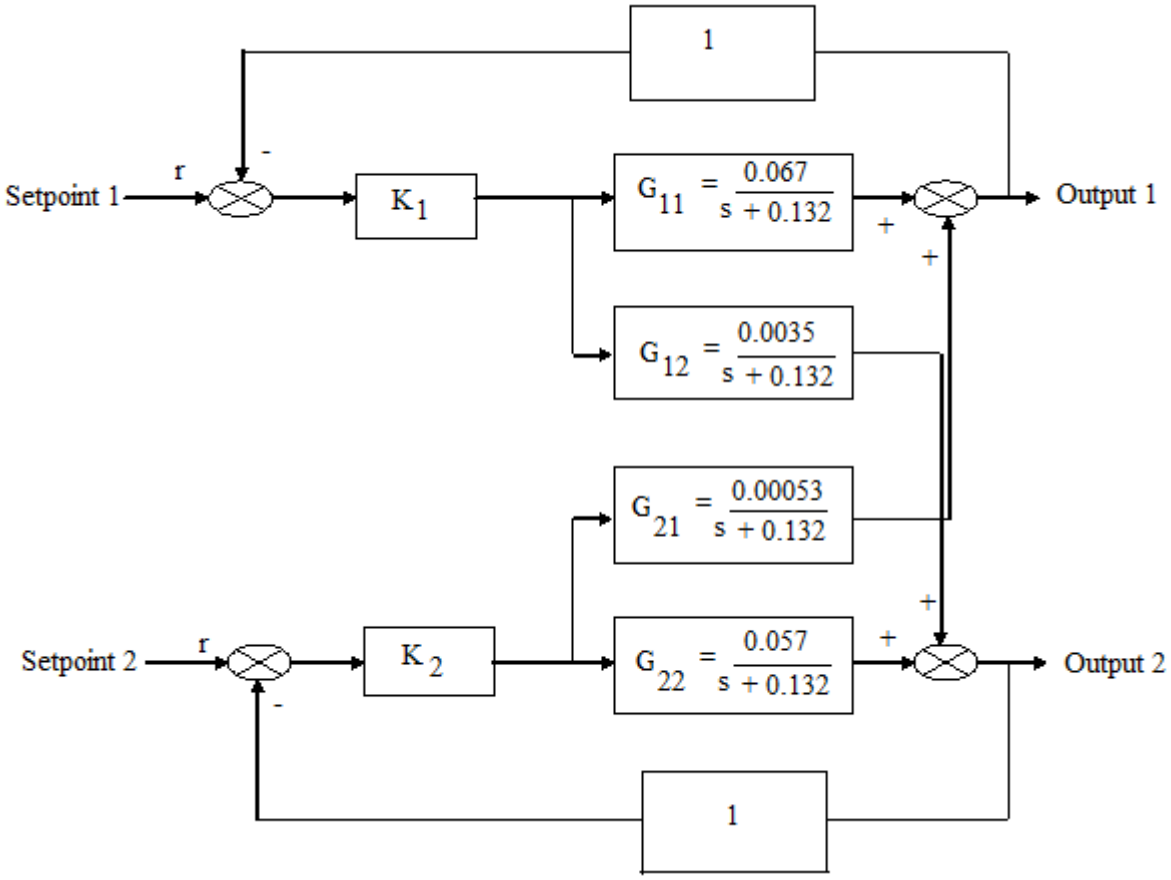

Figure 4.10: Diagram of the two feedback loops for a thermal system.

From this point the design of the thermal control is done by manipulating the input voltage for the heater blowers and the controlled output voltage that is the voltage due to the heat perceived by the sensors, which depends on its position relative to the heat blower position.

This information above, that's the basic building block of the thermal system, the input voltage which changes the temperature of the heater coils, and the sensed heat interpreted into the output voltage.

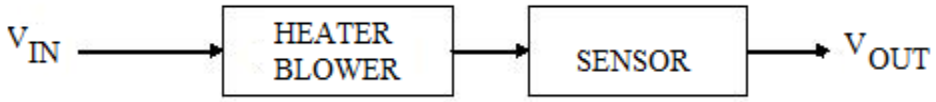

Figure 4.11: Block diagram showing the definition of the thermal system basic building block

4.6 Stability and Relative Stability

The two models given in the previous subsection can be seen to be stable as one has a pole at$s = -0.132$. The digitized model shows that it does not have any ringing.

The following figures below show the bode plots showing the stability margins of the continuous closed loop for G_{11} and G_{22} models respectively.

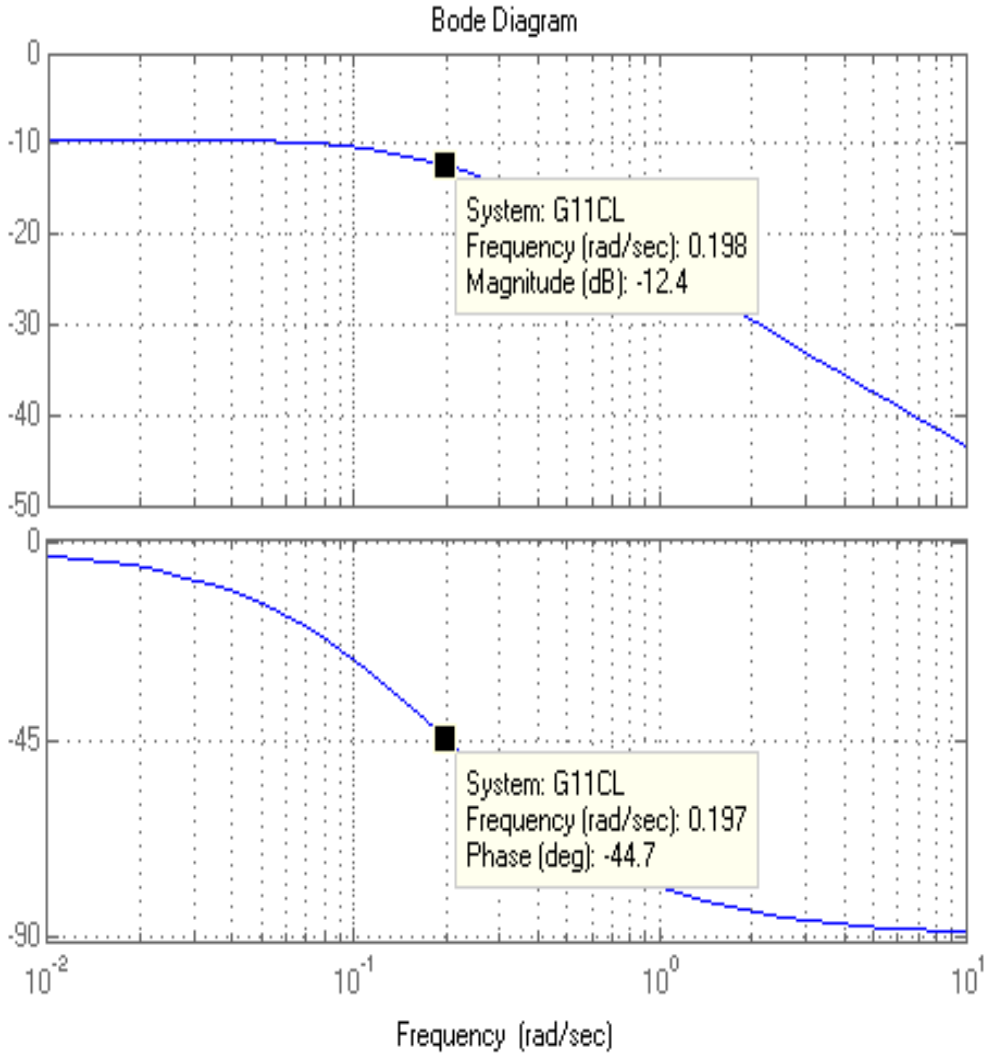

Figure 4.12: Bode plot for the continuous G_{11} closed loop model of the thermal system basic block showing the gain and phase margins within the plot

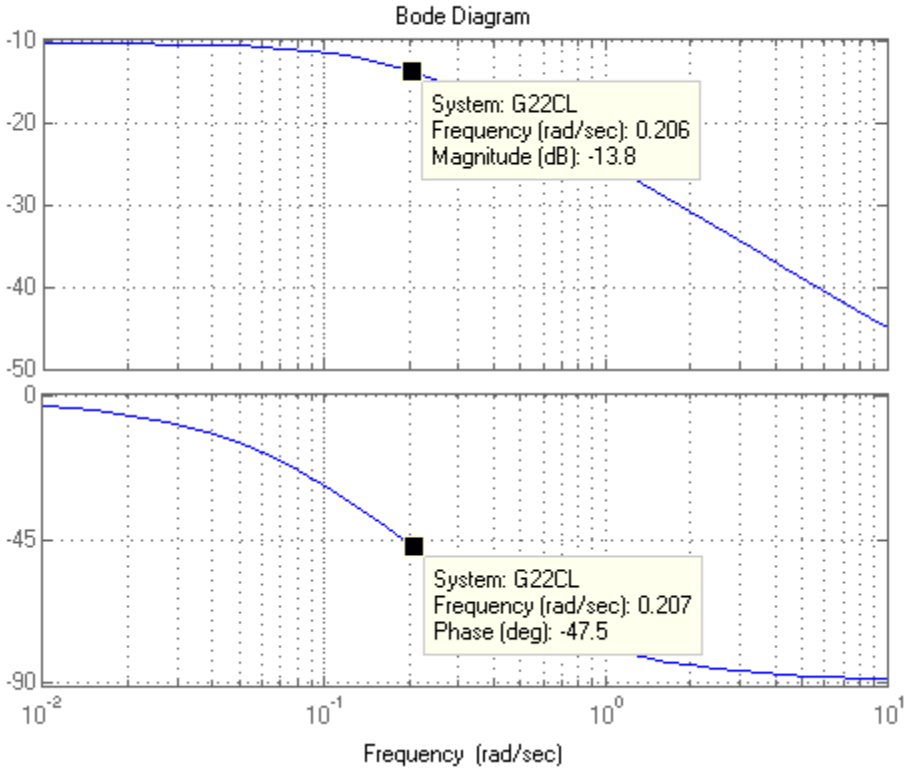

Figure 4.13: Bode plot for the continuous G_{22} closed loop model of the thermal system basic block showing the gain and phase margins within the plot

The transfer functions used however are approximates. These approximates were drawn by the fact that, the distance between the sensors and the corresponding heater blowers are approximately the same. Therefore, while all the time constants are approximately the same which gives the approximate diagonal transfer function matrix as,

$$G = \begin{bmatrix} g_{11} & 0 \\ 0 & g_{22} \end{bmatrix} = \begin{bmatrix} \frac{0.067}{s+0.132} & 0 \\ 0 & \frac{0.057}{s+0.132} \end{bmatrix} \quad 4.9$$

The author concluded that the full thermal system and the triangular structures are stable. This is because the two dominant models each has pole at $s = -0.132$ which is the left of the s – plane. The bode plots have the same gains for both magnitude plots and phase plot as shown in the bode plots. The stability shown by bode and root locus diagrams. This stable laboratory thermal system is good for the application of fault detection.

CHAPTER 5

PCA FAULT DETECTION TECHNIQUE AND APPLICATION ON THERMAL SYSTEM DATA

T^2 Statistics and the Q- statistics are the two methods that are applied when PCA is used for fault detection [39]. Each of these require that the optimal number of PCs to be retained is determined, as the T^2 Statistics uses the PCs that corresponds to the most significant eigenvalues while Q- statistics uses the PCs that corresponds to the less significant eigenvalues [56].

5.1 T^2 Statistics

In PCA for the thermal system data set $\boldsymbol{X}$ of n observations and m process variables, in this case the thermal data of two inputs and two outputs with 241 samples n = 241 and m = 4. After applying the singular value decomposition $\frac{1}{\sqrt{n-1}}\boldsymbol{X} = U\sum V^T$ which is equivalent to eigenvalue decomposition that gives the covariance matrix as;

$$\boldsymbol{S} = \frac{1}{n-1}\boldsymbol{X^T X} = \boldsymbol{V\Lambda V^T} \quad 5.1$$

Where the diagonal matrix $\Lambda = \sum^{\mathrm{T}}\sum \in R^{mxm}$ which are the eignevalues that are in magnitude of decreasing order [56].

The calculation of the T^2 Statistic values can help to detect a fault in a system. For the observation vector (a particular sample) x and $\boldsymbol{\Lambda} = \Sigma^T\Sigma$, where $\boldsymbol{\Lambda}$ is the eigenvalue, the value of T^2 Statistic becomes;

$$T^2 = x^T\boldsymbol{v}(\boldsymbol{\Sigma^T\Sigma})^{-1}\,\boldsymbol{v}^T x \quad 5.2$$

In the equation above v is the corresponding eigenvector of the significant eigenvalues which corresponds to the score space.

When the selected loading vectors a (correspond to the number of PCs) used for dimensional reduction of the original data presented as $\boldsymbol{P}$ matrix and $\boldsymbol{P} \in R^{mxa}$, then the T^2 statistics can be calculated as [42];

$$T^2 = \boldsymbol{x^T P}\sum_a^{-2}\boldsymbol{P^T x} \quad 5.3$$

Where $\mathbf{\Sigma_a}$ with the first a variables (the number of retained loading vectors) m rows and a columns of $\boldsymbol{\Sigma}$

5.1.1 T^2 Statistics Threshold

A threshold should be calculated to differentiate the techniques values that are within normal and those that are faulty. When the actual mean and the covariance matrix are known the threshold can be calculated using the following equation;

$$T_\alpha^2 = X_\alpha^2(a) \quad 5.4$$

Where X_α^2 is the percentage point of the X^2 distribution [Appendix C] but when only the sample covariance matrix is known the equation below is used;

$$T_\alpha^2 = \frac{(a(n-1)(n+1))}{(n(n-a))} F_\alpha(a, n-a) \quad 5.5$$

Where F_α is the percentage point of the F distribution [56, Appendix C].

5.2 Q – Statistics

The portion of the residual space which corresponds to the smaller or less significant eigenvalues can be monitored efficiently by the use of Q – statistic [5, 39, and 41].

$$Q = \boldsymbol{r^T r} \quad 5.6$$

Where $\boldsymbol{r} = (I - \boldsymbol{PP^T})x$ 5.7

5.2.1 Q – Statistics Threshold

The distribution of the Q – statistic that gives the threshold when using this technique of fault detection, can be approximated as;

$$Q_\alpha = \theta_1 \left[\frac{h_0 C_\alpha \sqrt{2\theta_2}}{\theta_1} + 1 + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1^2} \right]^{\frac{1}{h_0}} \quad 5.8$$

Where $\theta_i = \sum_{j=a+1}^{n} \delta^{2i}$ 5.9

And $h_0 = 1 - \frac{2\theta_1 \theta_3}{3\theta_2^2}$ 5.10

The C_α being the normal deviate corresponding to the $(1-\alpha)$ percentile (56) and $\delta^2 = \lambda$. The Q – statistic measures the random variations of the process while the threshold Q_α is applied to define the

normal variations of the random noise, and any violation of the threshold can indicate that the random noise has significantly changed [56].

Both of techniques can be used to detect different types of fault [39], in this dissertation only the T^2 Statistics will be used to find out the effect of the number of PCs retained in fault detection.

5.3 Fault Detection in Thermal system Using PCA

As mentioned above there two techniques that are used for fault detection in a process system [79]. Both of the techniques require the retention of the optimal number of PCs to be used in fault detection and these are explained in the succeeding sections.

When the process data is decomposed into principal components, the ones that correspond to the most significant eigenvalues that have the most information of the signal. The principal components that correspond to the insignificant eigenvalues usually have the noise that is in the signal.

T^2- statistic technique uses the principal components that have the most information of the signal to detect the fault [39]; this is the technique that is applied on the thermal data. The squared prediction error technique which is also known as the Q – statistic [56], which uses the residual part of the principal components was not applied in this research.

The laboratory heater blower thermal system data can be presented as matrix $\boldsymbol{X}$.

$$\boldsymbol{X} = \begin{bmatrix} u_1(t) & u_2(t) & y_1(t) & y_2(t) \\ u_{1,1}(t+1) & u_{2,1}(t+1) & y_{1,1}(t+1) & y_{2,1}(t+1) \\ \vdots & \vdots & \vdots & \vdots \\ u_{1,n}(t+n) & u_{2,n}(t+n) & y_{1,n}(t+n) & y_{2,n}(t+n) \end{bmatrix} \quad 5.11$$

With the variables input 1 ($u_1(t)$), input 2($u_2(t)$), output 1 ($y_1(t)$) and output 2 ($y_2(t)$).

The most important part of the PCA is the selection PCs that are corresponding to the largest eigenvalues to be retained [39]. This should be optimal in order to reduce the dimension of the original data but at the same time capture the variations of the data columns of the loading matrix $\boldsymbol{P} \in R^{mxa}$[56]. That's the ones that T^2 statistics technique uses for fault detection. This corresponds to the PCs associated with the first a eigenvalues; the projections of the observations in $\boldsymbol{X}$ into the lower dimensional space are contained in the score matrix,

$$\boldsymbol{T} = \boldsymbol{XP} \quad 5.12$$

T is projected into the m dimensional observation space.

$$\hat{\boldsymbol{x}} = \boldsymbol{TP^T} \quad 5.13$$

This gives the residual matrix

$$\breve{\boldsymbol{x}} = \boldsymbol{x} - \hat{\boldsymbol{x}} \quad 5.14$$

It is in the residual matrix where the variations in the observation space are put together in the loading vectors associated with the $m - a$ smallest singular values. That's the ones used in SPE technique (Q - Statistics). $\hat{\boldsymbol{x}}$ Spanned subspace is the score space it is in this space we have the process information. The $\breve{\boldsymbol{x}}$ spanned subspace is called the residual space and has the noise information of the signal [9].

5.4 T^2 Statistics on the Thermal System Process Data

On thermal process data this technique was applied on normal operation data, faulty data with three different magnitudes of data and faulty noise data. The faulty data with fault F1, the fault is introduced on the 100th to the 110th samples. This fault has the normal output 1 voltage increased above the normal by more than 100%, while fault F2 was introduced on the same number of samples on output 2 and the voltage was reduced by more than 100%. The final fault F3 was introduced on output 1 voltage on the same number of samples but the voltage for these samples was minimized by about 25%. The two faults, thus fault F1 and fault F2 could represent faults that can occur in the sensor while F3 is there for small spurious faults that may occur, like the change in environmental temperature.

5.4.1 T^2 Statistics on Normal data

This data was obtained in normal operation of the laboratory thermal system. Figure 5.1 shows the performance of this technique on this particular data. This figure also shows the 95% threshold in green [79]. The author chose this threshold to get most of the signal information. This varies with different number of PCs retained when calculating T^2 statistic values and the respective thresholds.

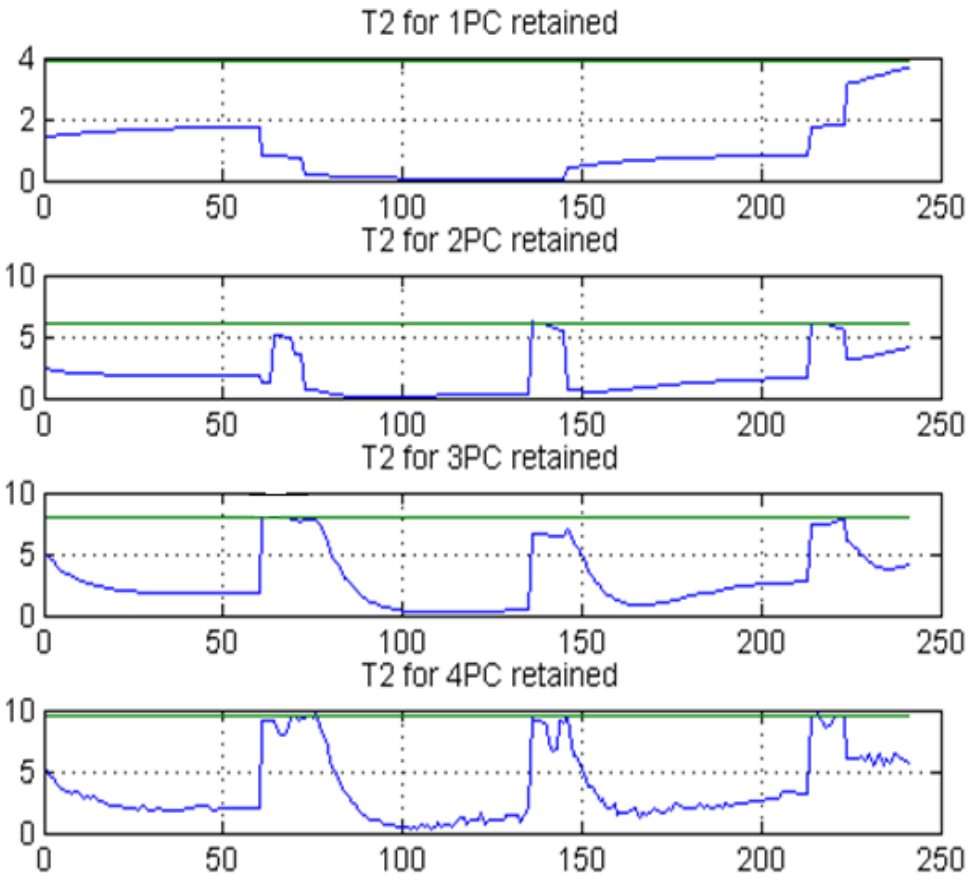

Figure 5.1: T^2 statistics and threshold plot for normal thermal process data.

5.4.2 T^2 Statistics on Fault F1 Faulty data

Three different faults were introduced on the system data on the same number of samples. The first fault F1, the size of the normal output 1 voltage was increased to almost double the size of the normal voltage. This was done on the 100^{th} to 110^{th} sample, to find out if this technique could pick the change with the retention of various numbers of PCs. The following figure shows the results of the experiment;

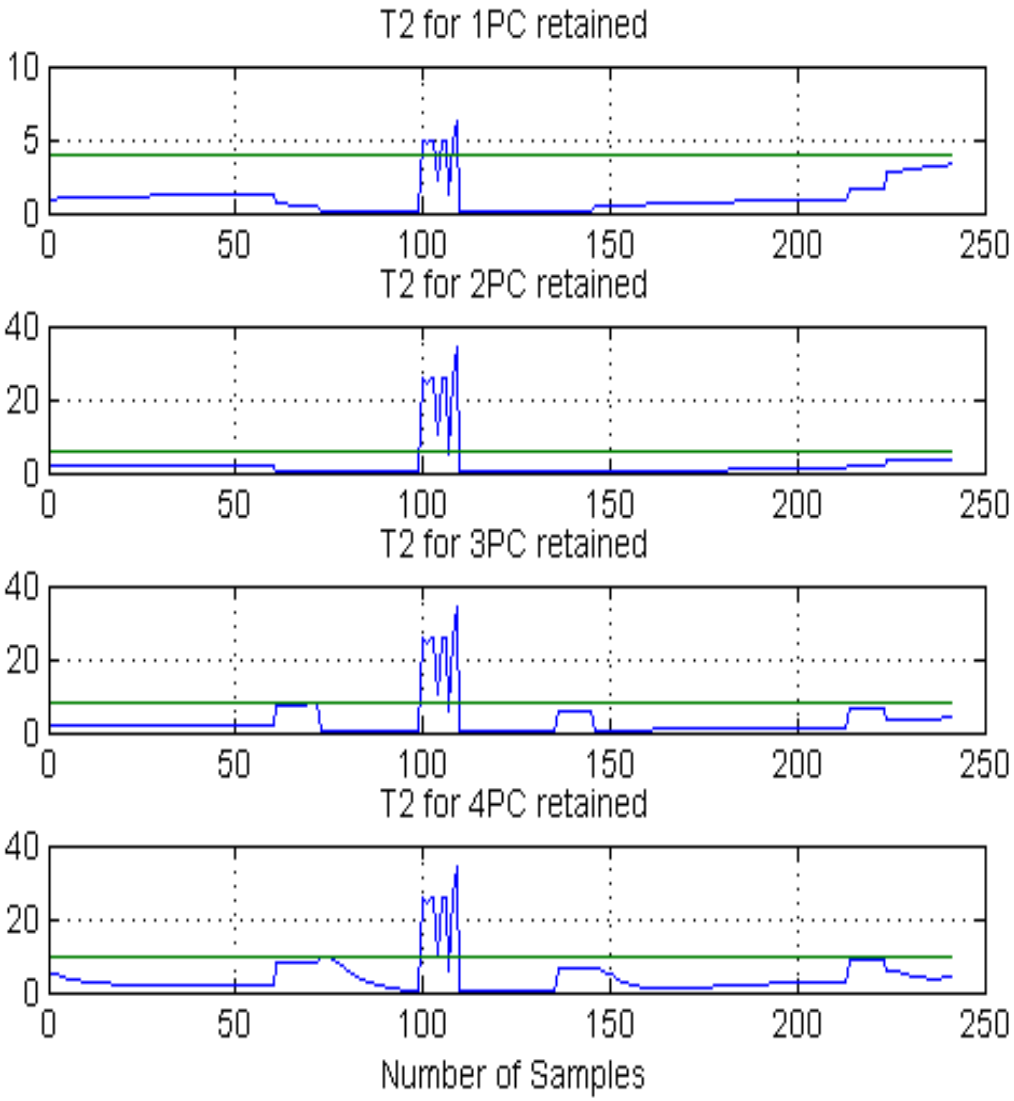

Figure 5.2: T^2 statistics and threshold plot for Faulty F1 in thermal process data.

The figure above shows that in each number of PC retention the fault F1 was detected. The detection was more prominent when two principal components were retained than any other number of PCs retained. One can see from the plot that for the threshold of 6.1 when two PCs were retained, almost the same values of T^2 statistics were obtained as for three and four PCs retained, T^2 statistics was able to detect fault F1.

5.4.3 T^2 Statistics on Fault F2 Faulty data

The author realized the importance of noticing the small changes that may have effect on the control system. Fault F2 fault was introduced in the process data, which has the lower voltage compared to the normal voltage. As for fault F1, fault F2 fault was introduced on the 100th sample to 110th samples. The figure 5.3 shows the plot of the fault detection with the PCs retention at different times. It is shown that with one PC retained this particular fault could not be detected but the rest number of PCs could detect the fault. This means that the information that was retained in the signal with the retention of one PC was not enough to pick up fault F2. The reduction of the data dimension for this one PC accomplished but the fault F2 was on the residual space discarded in the three remaining PCs.

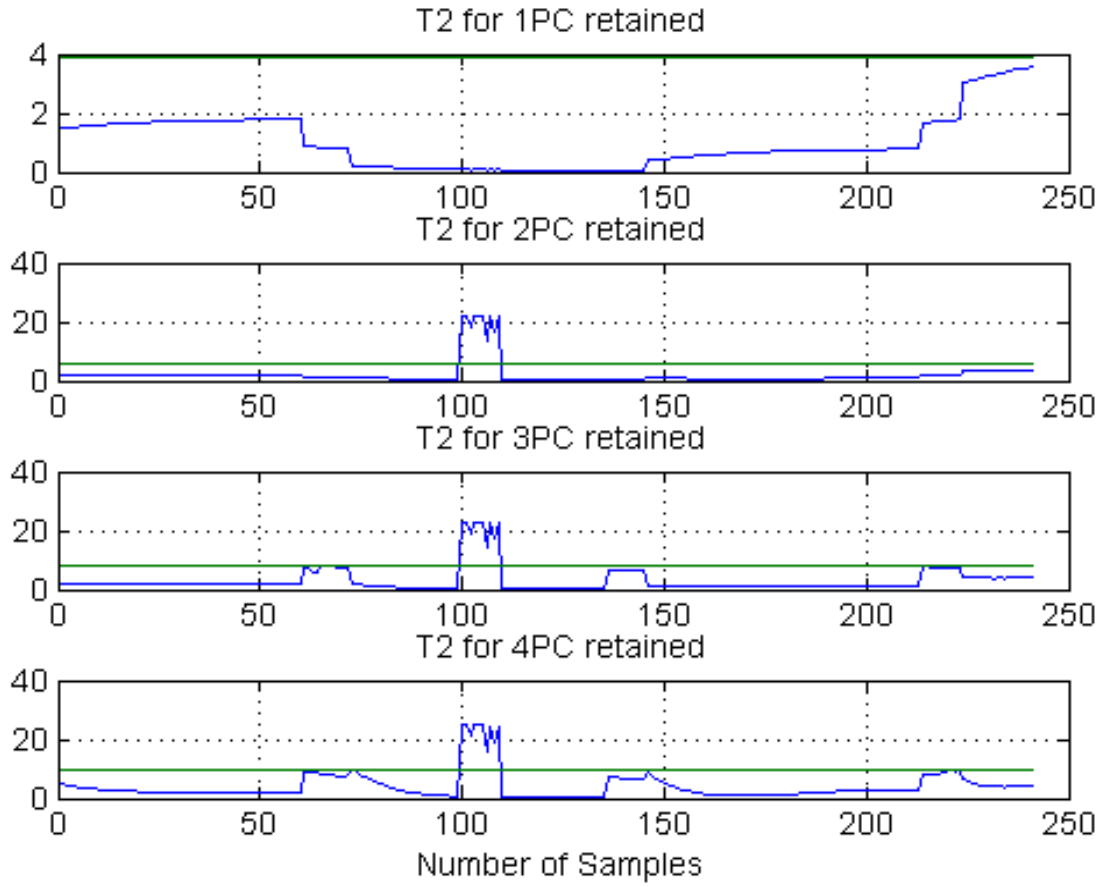

Figure 5.3: T^2 statistics and threshold plot for Faulty F2 thermal process data.

The values of T^2 statistics for two and three PCs retained is almost the same but slightly different for four PCs retained which is a bit higher. But PCA is a dimensional reduction technique and for this idea to be satisfied at the same time be able to detect F2, the retention of two PCs is the most appropriate.

5.4.4 T^2 Statistics on Fault F3 Faulty data

Fault F3 was introduced to detect faults that occur and cause a very slight change in the normal output voltage. The difference in voltage is very small; such types of faults could be affected by the change in humidity of the environmental air. This fault was introduced on the same number of samples, the 100th to 110th samples in the thermal process data and figure 5.4 has the plot of the fault detection with the different number of PCs retained.

It can be noticed from the plot below that the fault F3 could only be detected with the retention of four PCs. The fault could not be detected when one, two or three PCs are retained. This means no reduction of the dimension data is not achieve and shows that the fault F3 appeared in the noise part of the system signal hence the detection by retaining all the four PCs.

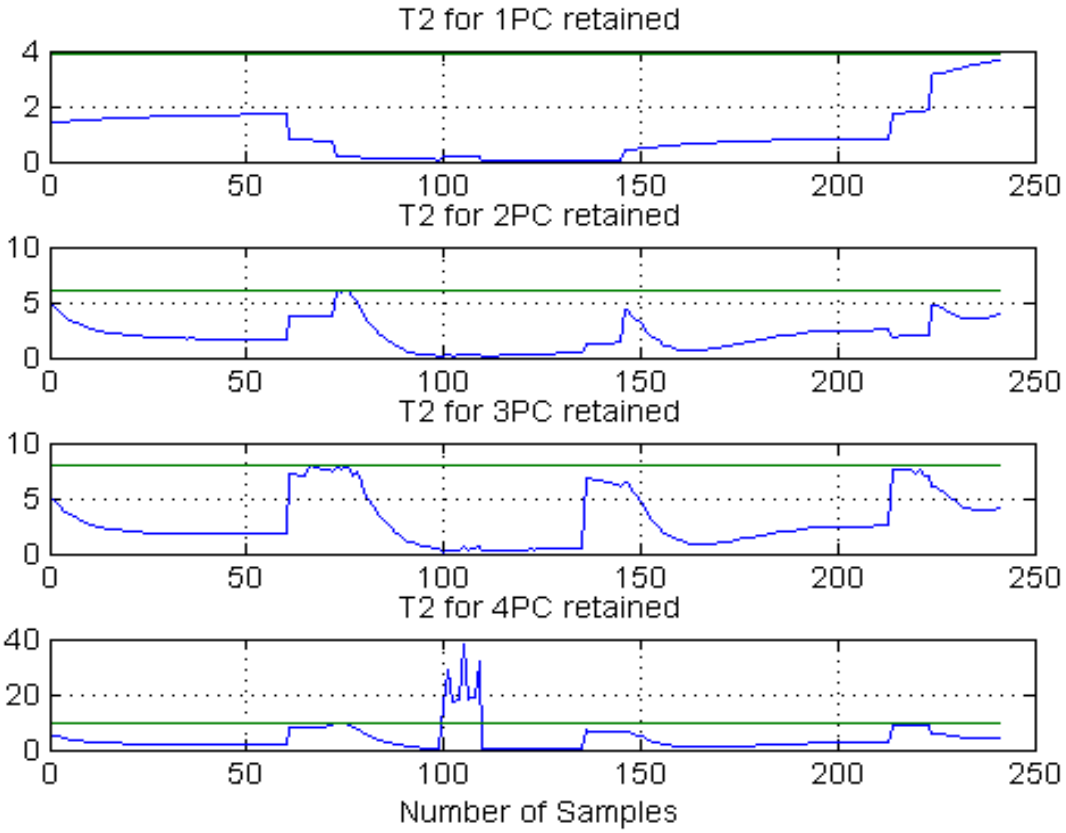

Figure 5.4: T^2 statistics and threshold plot for Faulty F3 thermal process data.

The same procedure of fault detection was done on the process data where some noise of about 0.5dB was introduced. The purpose was to find out the effects of noise on the T^2 technique of PCA fault detection. Value of the noise introduced was supposed to be significant but not too much completely compromise

the output voltage and at the same time not too small not to have an effect on the output voltage. This is how the choice of the noise of 0.5dB was done and introduced on the thermal system day with the three faults to observe the effect with the fault detection.

5.5 Faulty Noise Data

It is important to see effect of noise on the values of T^2 statistics. The noise of 0.5dB was introduced in the system and the three faults as well. This was selected as in reality have an effect on the output voltage of the system but not too much or too small to have an insignificant effect. The following are the results of the hotellings T^2 statistics values from the noisy data with different number of PCs retained with the different faults.

5.5.1 T^2 Statistics for Fault F1 Faulty Noise Data

The F1 fault represents high output voltage (voltage above the normal output voltage value) on the noise data. It can be noted that the fault was detected by all the PCs retained, but more prominent when two, three and four PCs were retained. All these had almost the same values of T^2 statistics at the point of fault occurrence as can be observed from the plot in figure 5.5.

The effect of the noise on the faulty data as observed on the figure 5.5 following, the sensitivity is high for this fault F1 as compared to the faulty data without noise. The T^2 statistics was able to detect the fault F1 with much sensitivity that is shown by the higher values of T^2 statistics plotted.

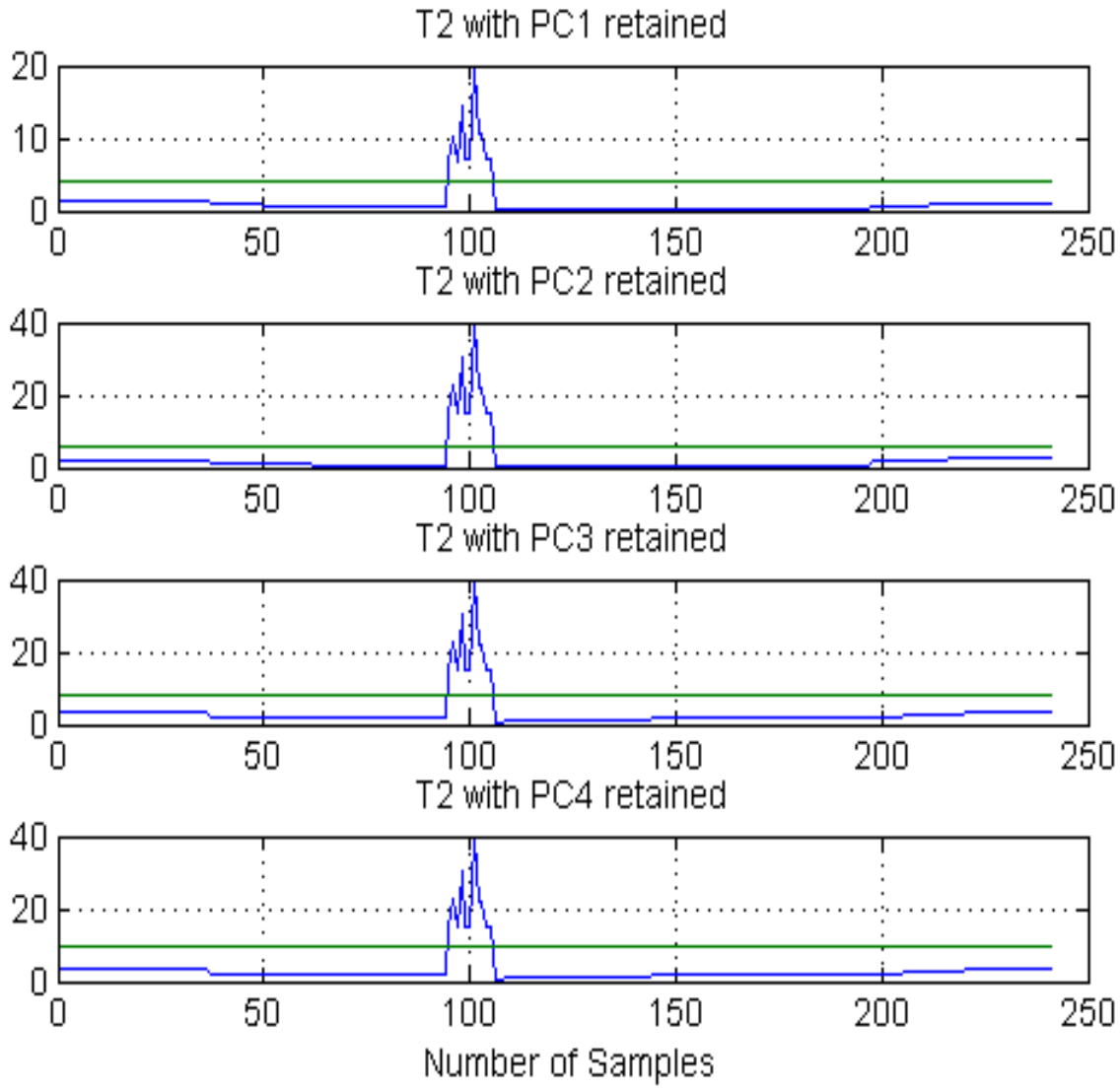

Figure 5.5: T^2 statistics and threshold plot for Faulty F1 thermal process data.

5.5.2 T^2 Statistics for Fault F2 Faulty Noise Data

In figure 5.6 fault F2 could not be detected with the retention of one PC but the rest PCs retained detected the fault with almost the same values as the fault F2 without noise with a slight difference. This time T^2 statistics was a bit more sensitive to this fault F2 with the introduced noise than the one without the introduced noise in it.

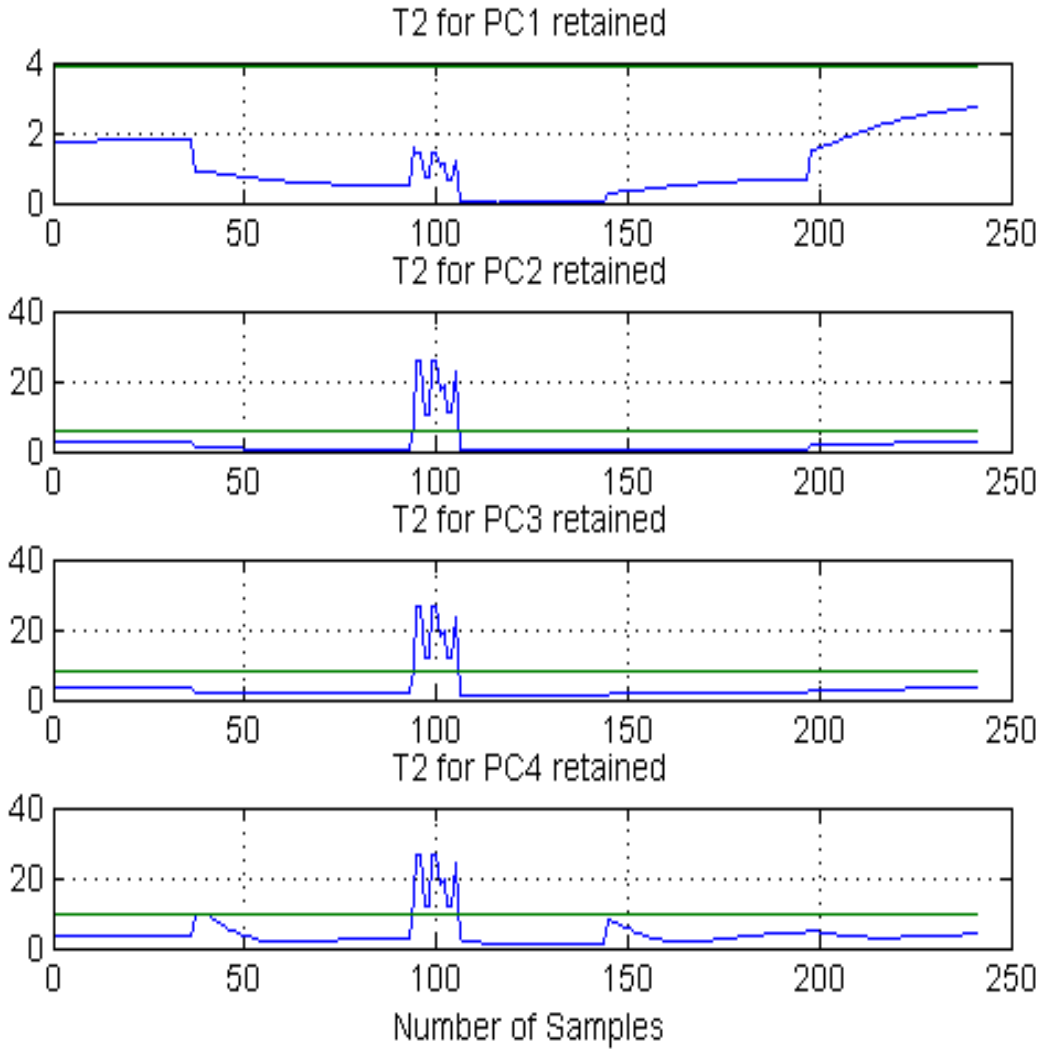

Figure 5.6: T^2 statistics and threshold plot for Faulty F2 thermal process data with 0.5dB noise.

5.5.3 T^2 Statistics for Fault F3 Faulty Noise Data

The retention of one PC and two PCs on the fault F3 of the noisy data, could not detect the fault as observed in figure 5.7. The fault F3 could only be picked by the retention of three and four PCs with almost the same values. The introduction of the 0.5dB noise has made the T^2 statistics values be sensitive to fault F3 with the retention of not only four PCs but also with three PCs. The introduction of the noise spread the fault on the wider residual space.

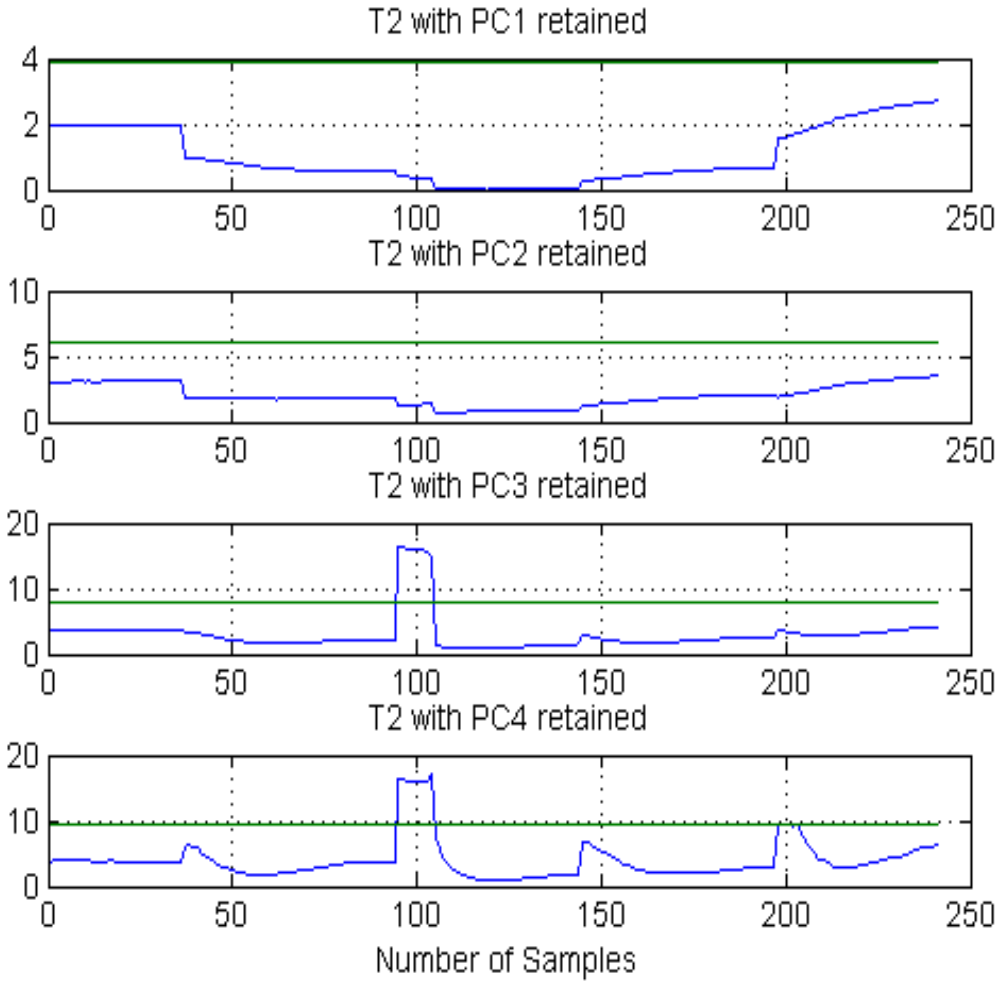

Figure 5.7: T^2 statistics and threshold plot for Faulty F3 thermal process data.

An experiment was run to see the effect of retaining different numbers of PCs on the three faults with fault voltages varied and observe the performance. The faulty output voltages were varied but still keeping the main idea of each fault. Table 5.1 below has the results.

Fault Detection Table

	F1				F2				F3			
F1	PC_1	PC_2	PC_3	PC_4	PC_1	PC_2	PC_3	PC_4	PC_1	PC_2	PC_3	PC_4
10V	Yes	Yes	Yes	Yes								
8V	Yes	Yes	Yes	Yes								
6V	Yes	Yes	Yes	Yes								
5V	No	Yes	Yes	Yes								
F2												
0V					No	Yes	Yes	Yes				
1V					No	Yes	Yes	Yes				
2V					No	Yes	Yes	Yes				
3V					No	Yes	Yes	Yes				
F3												
3.5V									No	No	No	Yes
4V									No	No	No	Yes
4.2V									No	No	No	Yes

Table 5.1: Fault detection table

Fault Detection Table

	F1				F2				F3			
F1	PC1	PC2	PC3	PC4	PC1	PC2	PC3	PC4	PC1	PC2	PC3	PC4
0.1dB	Yes	Yes	Yes	Yes								
0.2dB	Yes	Yes	Yes	Yes								
0.3dB	Yes	Yes	Yes	Yes								
0.5dB	Yes	Yes	Yes	Yes								
F2												
0.1dB					No	Yes	Yes	Yes				
0.2dB					No	Yes	Yes	Yes				
0.3dB					No	Yes	Yes	Yes				
0.5dB					Yes	Yes	Yes	Yes				
F3												
0.1dB									No	No	No	Yes
0.3dB									No	No	No	Yes
0.5dB									No	No	No	Yes

Table 5.2: Fault detection table with varied noise and fault detection.

T^2 statistic technique is able to detect all the three different types of faults when PCA is used for fault detection. The different number of PCs retained had an effect on fault detection when using this technique, as observed some number of PCs retained could not detect some types of faults. F3 could only

be detected when all the PCs were retained and no dimension reduction could be done for this type of fault. This shows the importance of considering the number PCs to be retained, while reducing the dimension of the process data, when using PCA as a method of fault detection.

Table 5.1 3.5V fault voltage is almost 20% closer to the normal output voltage for fault F3 and has the characters of fault F2 as far as fault detection is concerned, this shows that this fault does need to be considered for this fault.

Fault F1 and fault F2, the retention of two principal components looks good and suitable for both fault detection and dimension reduction satisfying the use of PCA method [5]. Fault F3 could only be detected by retaining four which are all the PCs and the purpose of dimension reduction is not fulfilled for PCA. This is could be an advantage because the thermal system will not recognize any change in environmental temperature or humidity as a fault, when the fault detection setting is set to retain less that four PCs which is the main purpose of using PCA.

The introduction of a small value of noise was insignificant on the output voltage. According to table 5.2 0.5dB of noise increased the sensitivity of the fault especially F3 which could be detected by the retention of three PCs and not only four PCs as when the small amount of noise is introduced, which shows the noise increase the residual space.

CHAPTER 6

CONSIDERED METHODS FOR DETERMINING OPTIMAL NUMBER OF PRINCIPAL COMPONENTS TO BE RETAINED FOR FAULT DETECTION

There are more than ten methods that are used by researchers used to determine the optimal number of PCs to be retained for effective fault detection [39]. These methods are applied with the aim of effectively detect the fault at the same time as dimension reduction for the system data.

Only the most commonly used methods, the Scree test method and Cumulative Percent Variance (CPV) [39], are to be compared with relatively recent signal to noise ratio method (SNR) of number of PCs determination in this chapter.

Most of these methods give just one particular number of PCs to be retained for the specific system data with different faults. It is important to know what number of PCs is sensitive to which faults for the fault detection technique to be effective to save the system and the operation personnel.

6.1 The Scree Test method

This is one of the most commonly used methods which involve plotting the eigenvalues against the number of principal components. These eigenvalues are arranged in ascending [38] and linked with a line. This method states that the number of PCs to be retained is the value on the slope of the line joining the plotted points called the 'steep' [38]. This value should be to the left of this 'steep' and not to the right. Figure 6.1 gives the number of PCs to be retained for the normal data, F1, F2 and F3.

It can be observed that for normal data, F2 and F3, two PCs are determined to by this method to be retained, to both reduce the dimension of the system data and for fault detection.

For F1 three PCs that are determined to be retained for fault detection and reduction of the system order. All this is shown in the following figure, the technique of fault detection is applied with these determined number of PCs by the Scree method and the result are in figure 6.2.

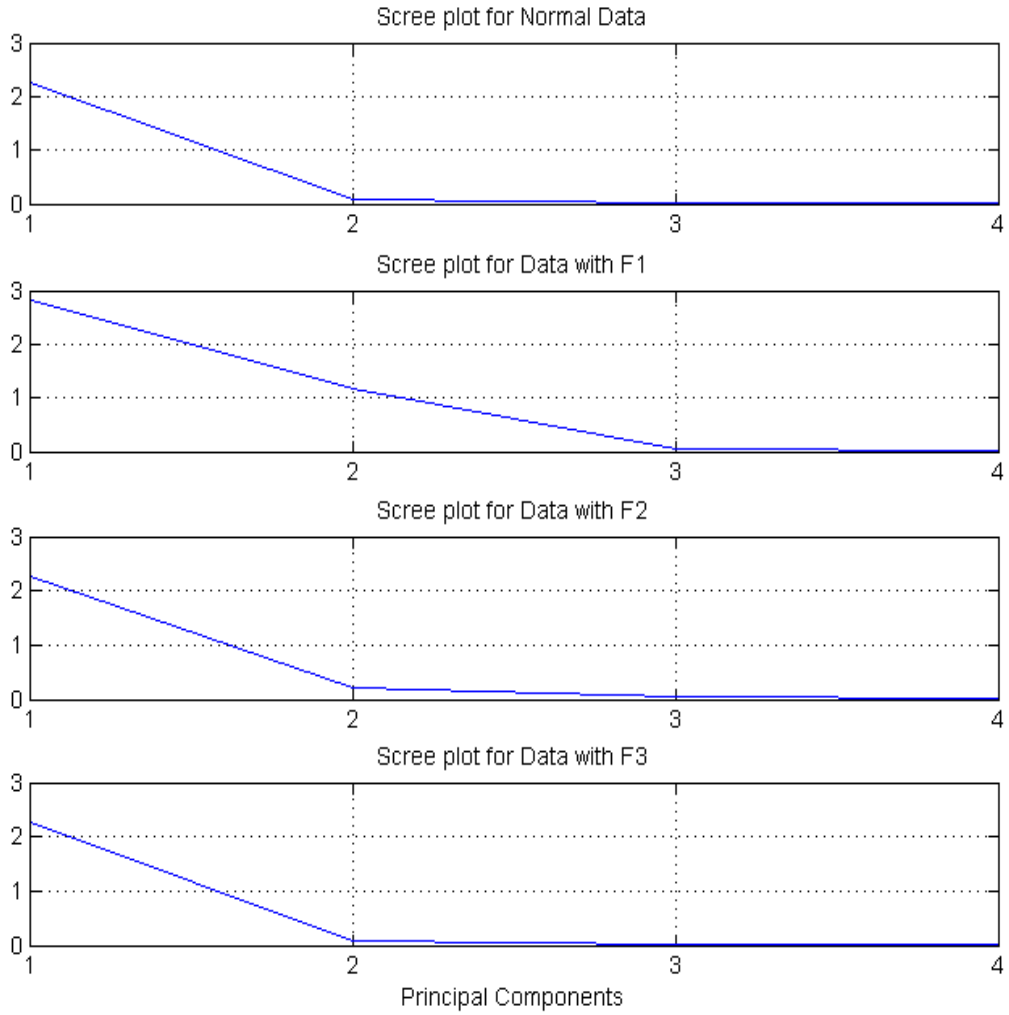

Figure 6.1: Scree test plots for normal data, F1, F2 and F3 data.

The following figure shows the T^2 statistics values for the three faults with the determined PCs by the Scree test method. The results are shown in figure 6.2.

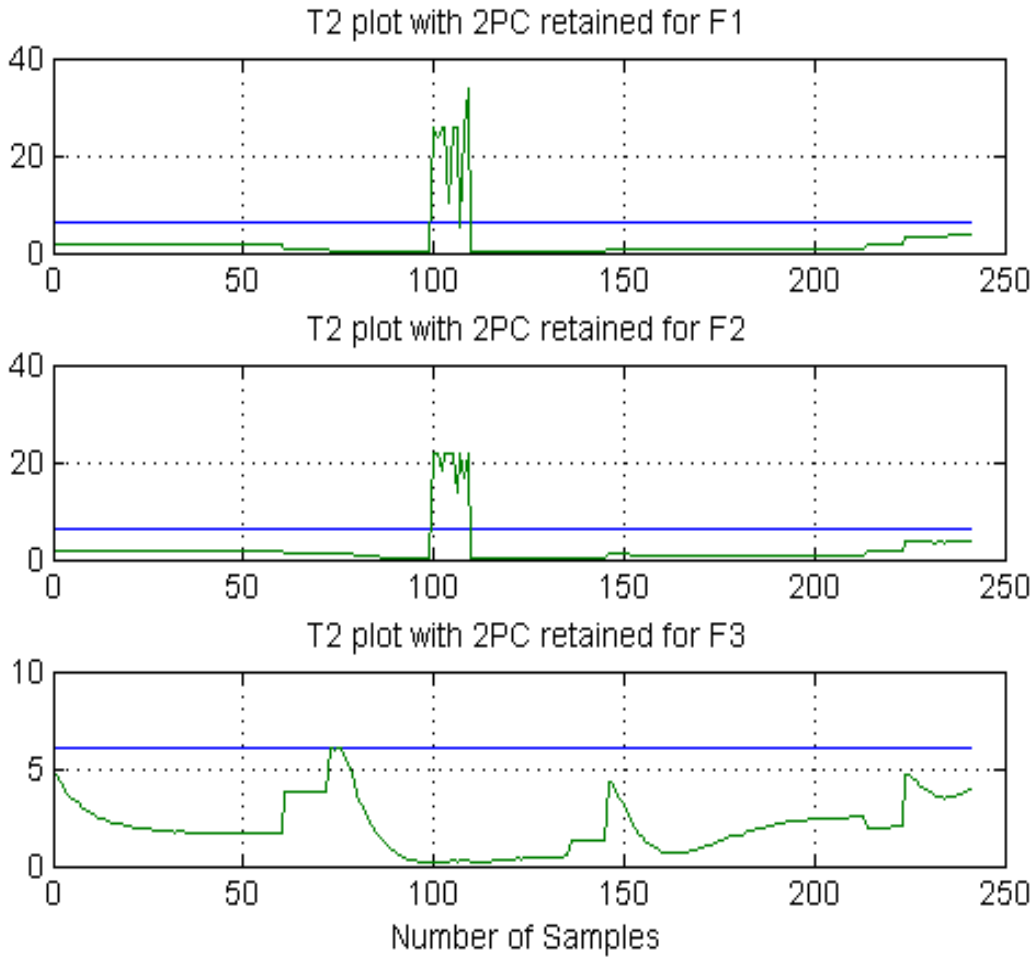

Figure 6.2: T^2 statistics values for F1, F2 and F3 with two PCs retained.

F3 could not be detected with the two PCs that were retained after being determined by the scree method. This shows that this method is not right all the times when it comes to determining the optimal number of PCs to be retained for fault detection when using PCA.

6.2 The Cumulative Variance Percent Method

Cumulative variance percent method is just one of the many methods used for determining the number of PCs to be retained when using PCA [39].

This method selects the percentage of the total variation which is desired for the selected PCs. [38] Shows that one is at liberty to choose the percentage one is comfortable with, this could be 80%, 95% or 99% of the total variation.

$$CPV = 100\left[\frac{\sum_j^l \lambda_j}{\sum_j^m \lambda_j}\right]\% \qquad 6.1$$

The following figure has the cumulative variance percentage for the normal data, the data with fault 1, data with fault 2 and data with fault 3.

The author considered 95% to be the cumulative variance for determining the number of PCs to be retained for dimension reduction for the normal data, the choice was done to get the most of the information of the signal in the score space. The same CPV was used for PC determination for both dimension reduction and fault detection for the data with faults.

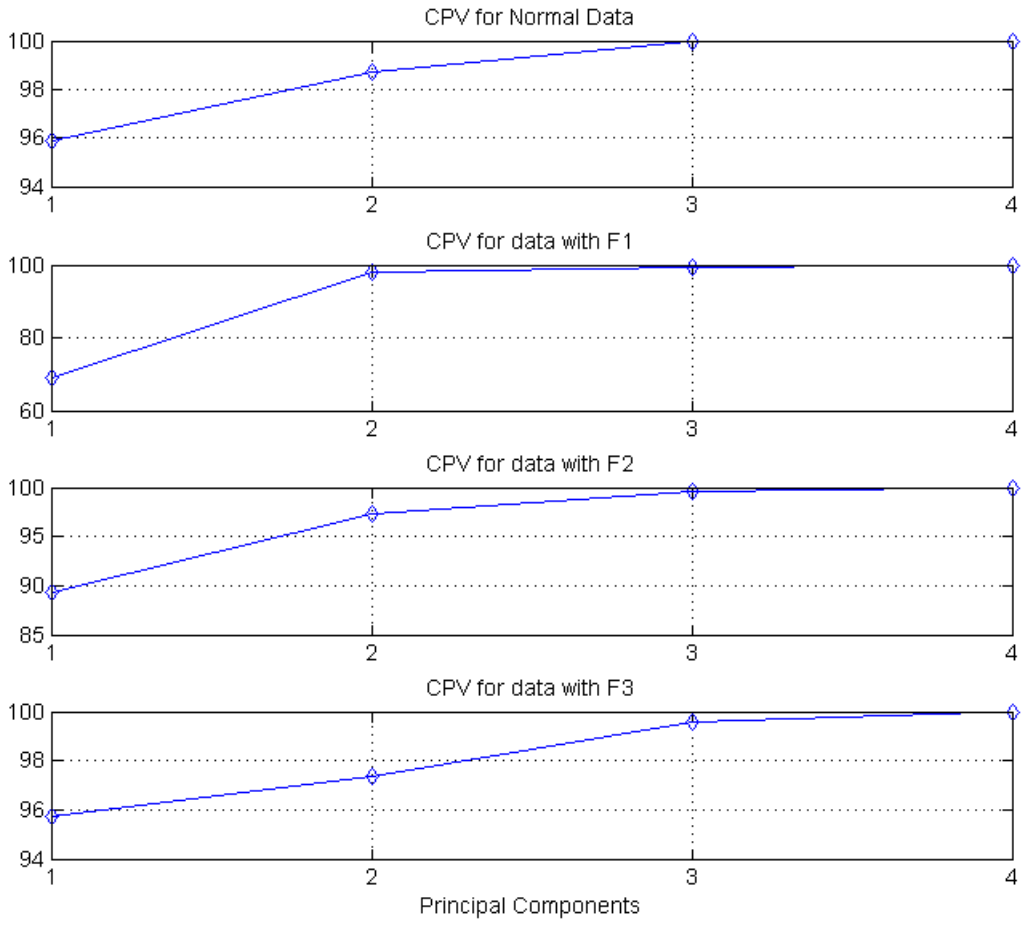

Figure 6.3: CPV plot for Normal data and the data with faults

It is observed from the plot above that with the 95% CPV; it determined one PC for normal data and F3, two PCs for F1 and F2.

These numbers of the retained PCs were used in fault detection with the calculation of T^2 statistics values which are plotted in the following figure.

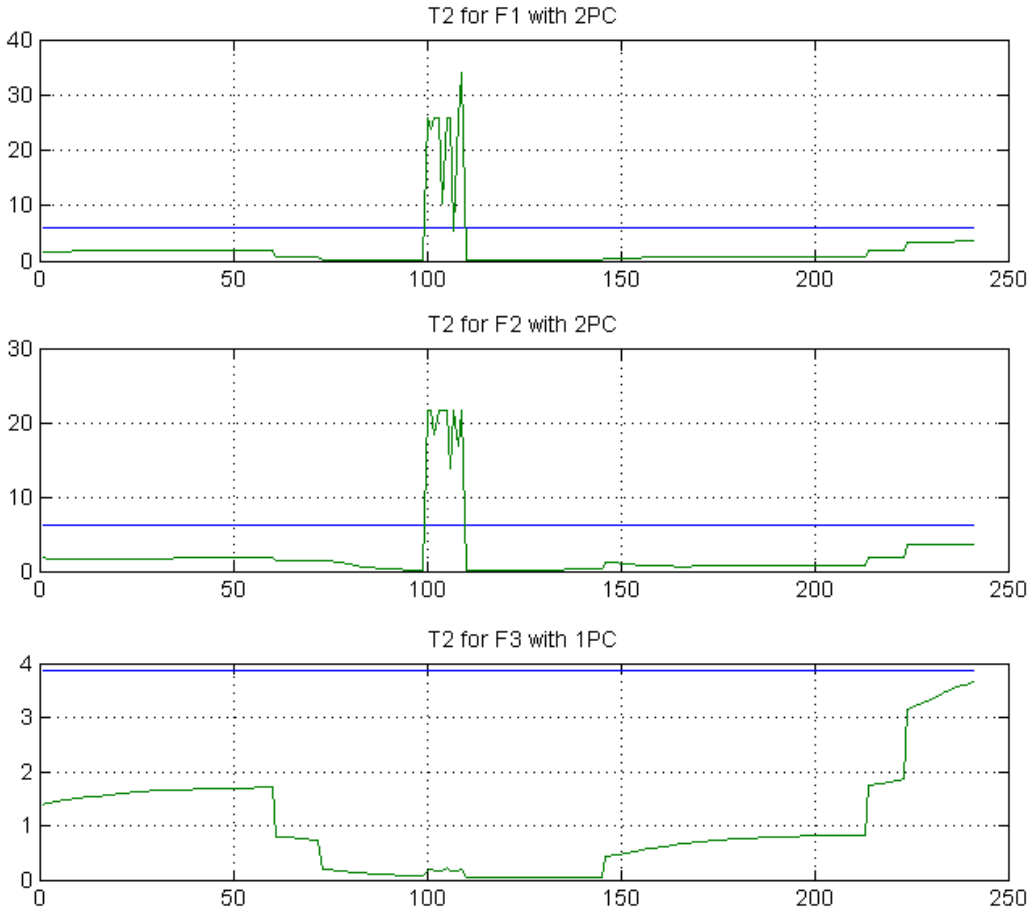

Figure 6.4: T^2 Statistics for all samples for F1, F2 and F3

Fault 3 could not be detected with the retention of one PC which was determined by CPV method. It worked for the dimension reduction but not feasible for fault detection. This method worked for fault 1 and fault 2, where the fault was detected with the T^2 statistics technique.

6.3 The Fault Signal to Noise Ratio Method

This is relatively the most recent method of determining the number of PCs to be retained focusing mainly on fault detection. This method can only be applied to data with priori information and the results gives which number of PCs is more sensitive to which particular fault [39].

The following figure has the T^2 statistics values plotted with various numbers of PCs determined for the three different faults. Figure 6.5(a) has the plot for F1, while the next has that with fault 1 with a noisy data.

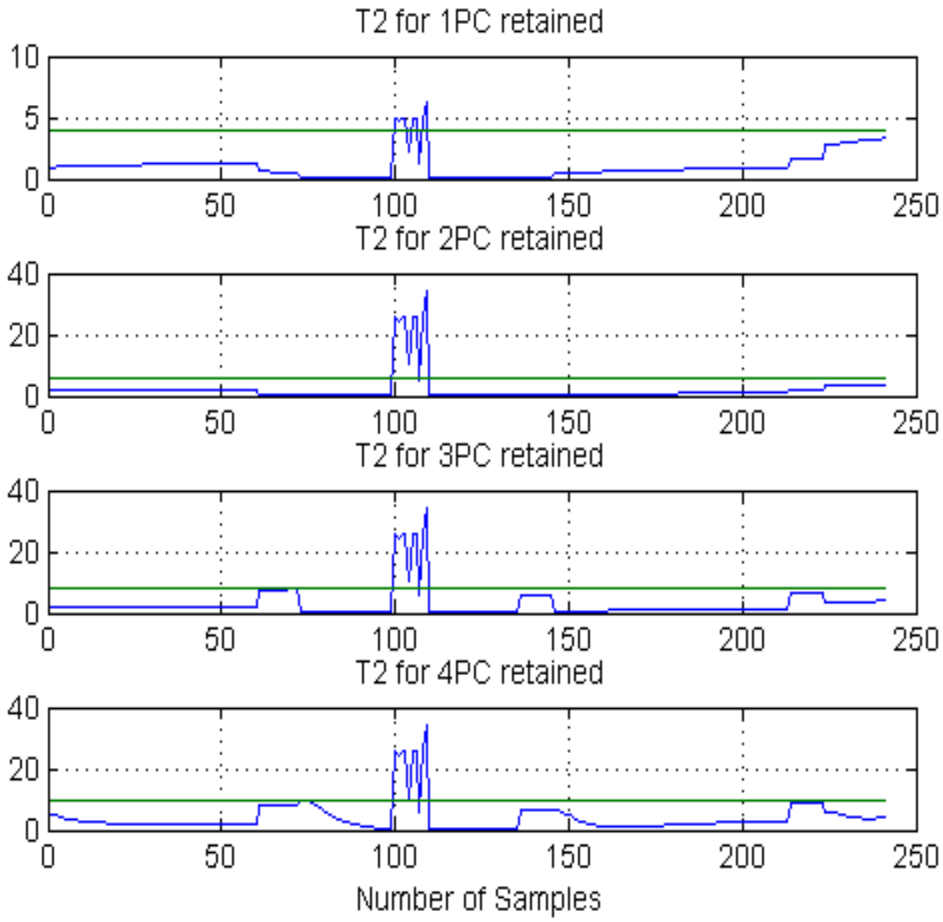

Figure 6.5: T^2 statistics plot for F1 of the thermal data for various PCs retained.

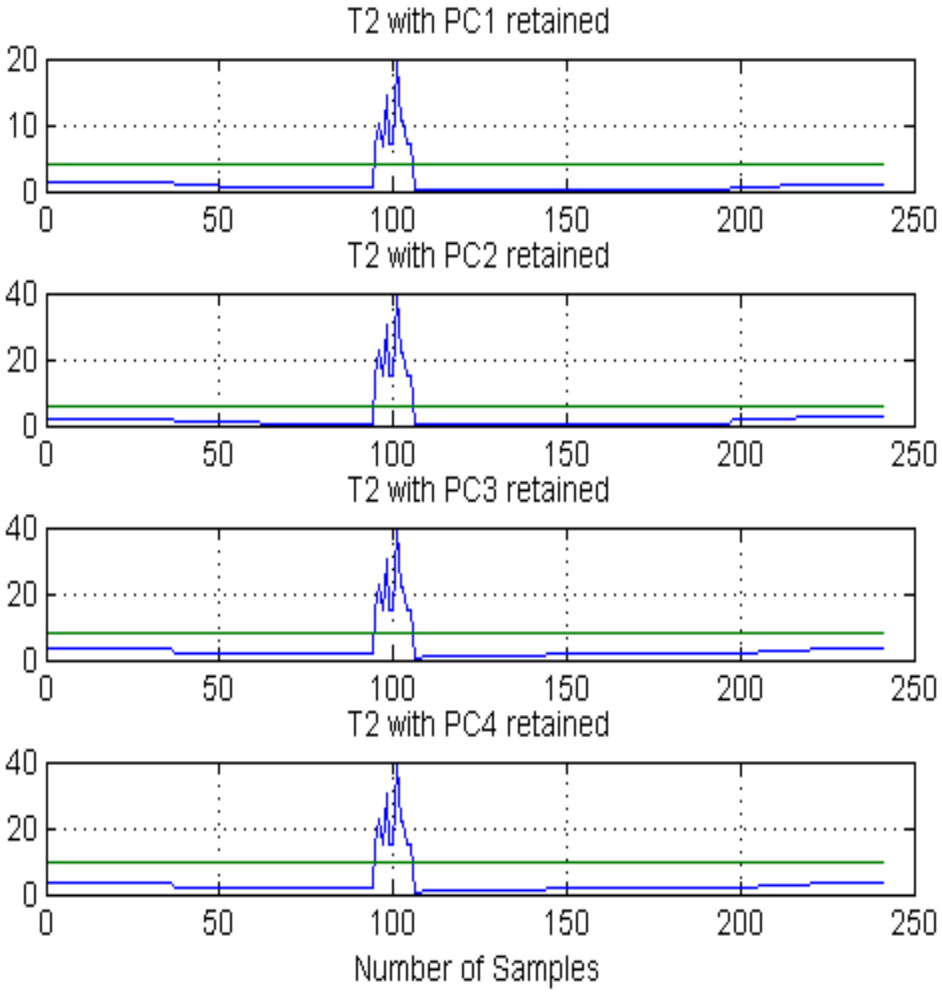

Figure 6.6: T^2 statistics plot for F1 of the thermal data with 0.5dB noise for various PCs retained.

It can be seen that in both data plots fault 1 could be detected with the retention of either four PCs. One has to consider the sensitivity of the fault with a particular number of PCs retained and at the same time consider the dimension reduction of the data. In this case the most optimal number of PCs to be retained should be two for fault 1, with the high sensitivity for fault detection as for there and four PCs.

Figure 6.6 has the plot of the laboratory thermal system data with fault 2, and the same fault in a noisy data.

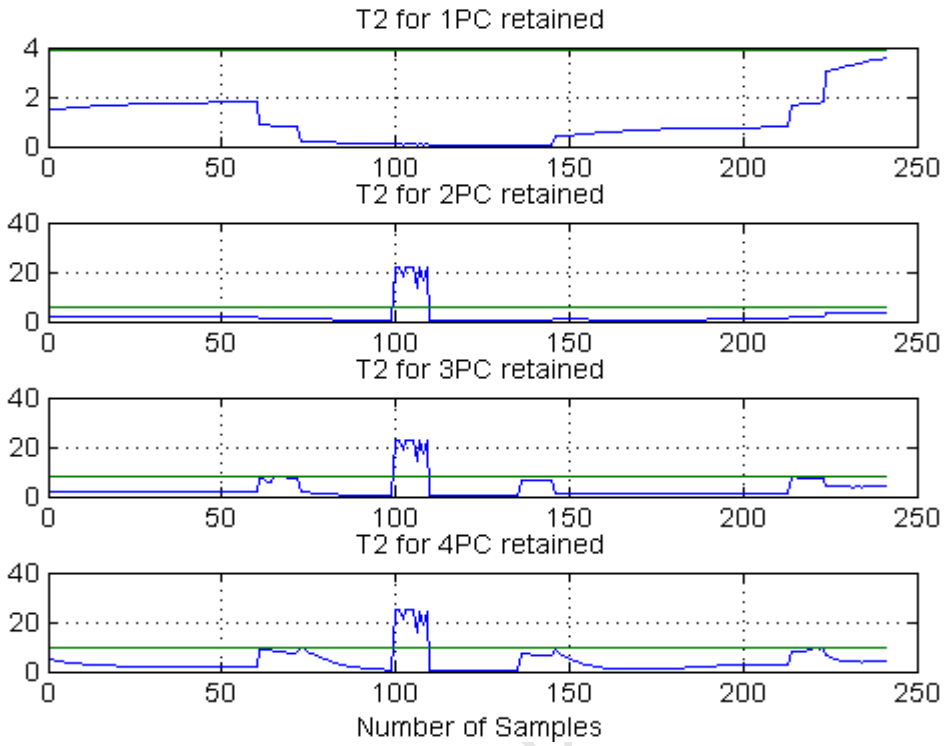

Figure 6.7: T^2 statistics plot for F2 of the thermal data for various PCs retained.

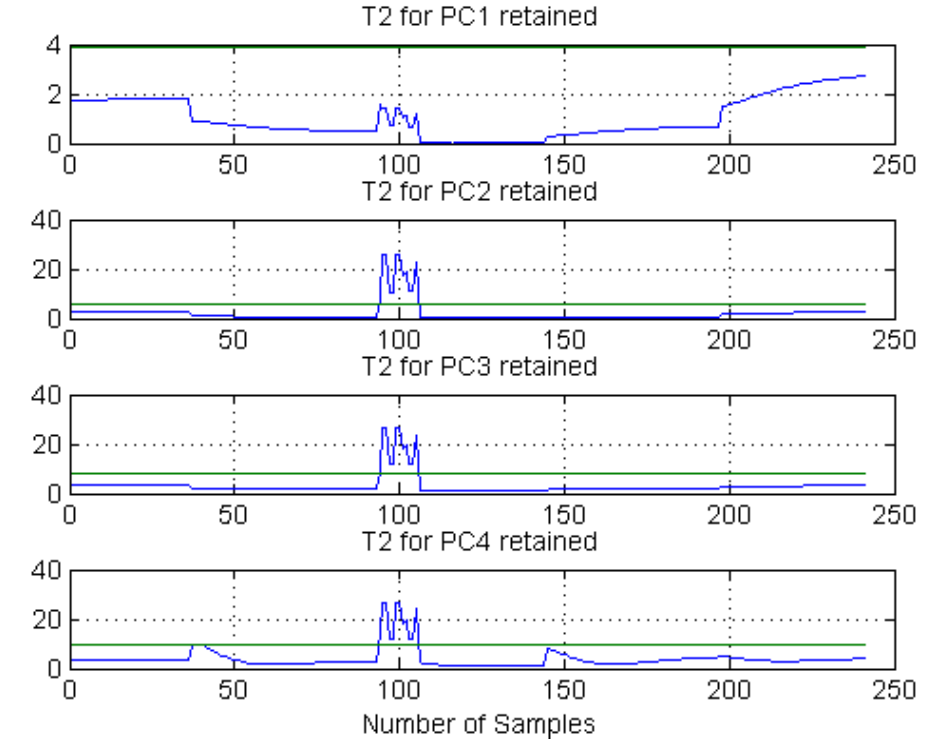

Figure 6.8: T^2 statistics plot for F2 of the thermal noisy data for various PCs retained.

Fault 2 could not be detected by retaining one PC, but with the rest of the PCs. The sensitivity of the fault 2 detection is almost the same for the three combinations of the PCs that could detect the fault, making it more plausible to consider two PCs as the optimal.

The same procedure was done for fault 3 and the results given in the following figures.

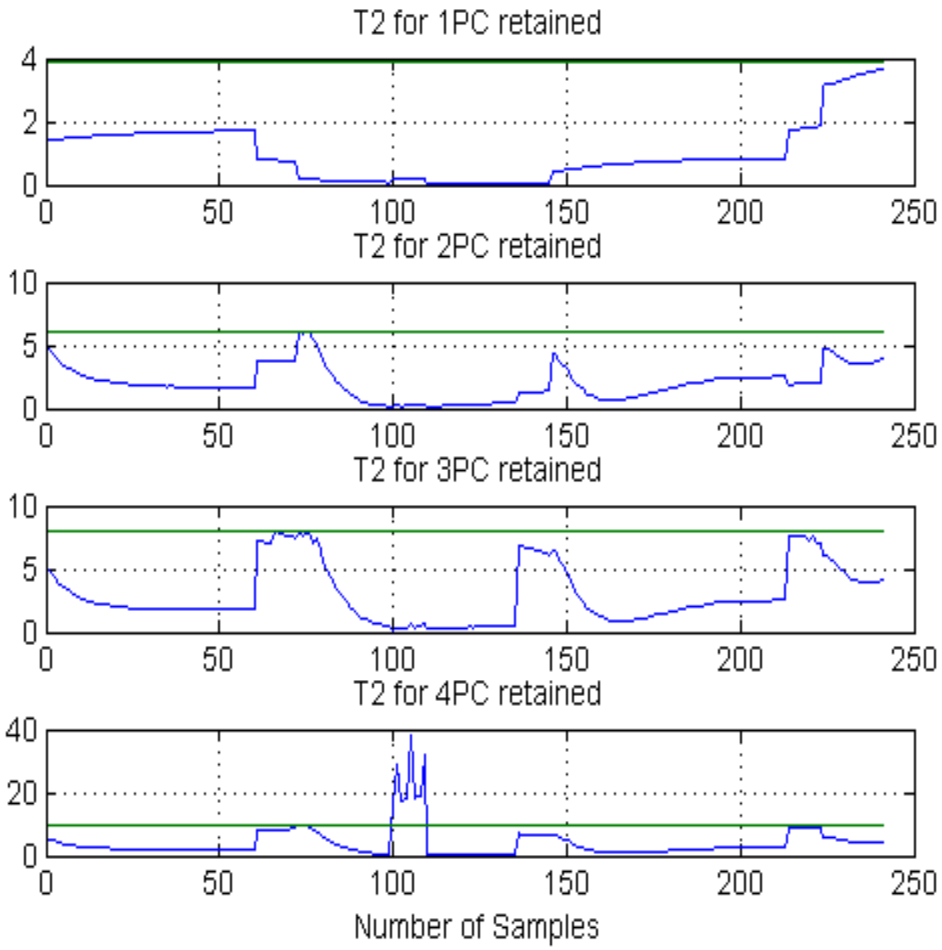

Figure 6.9: T^2 statistics plot for F3 of the thermal data for various PCs retained.

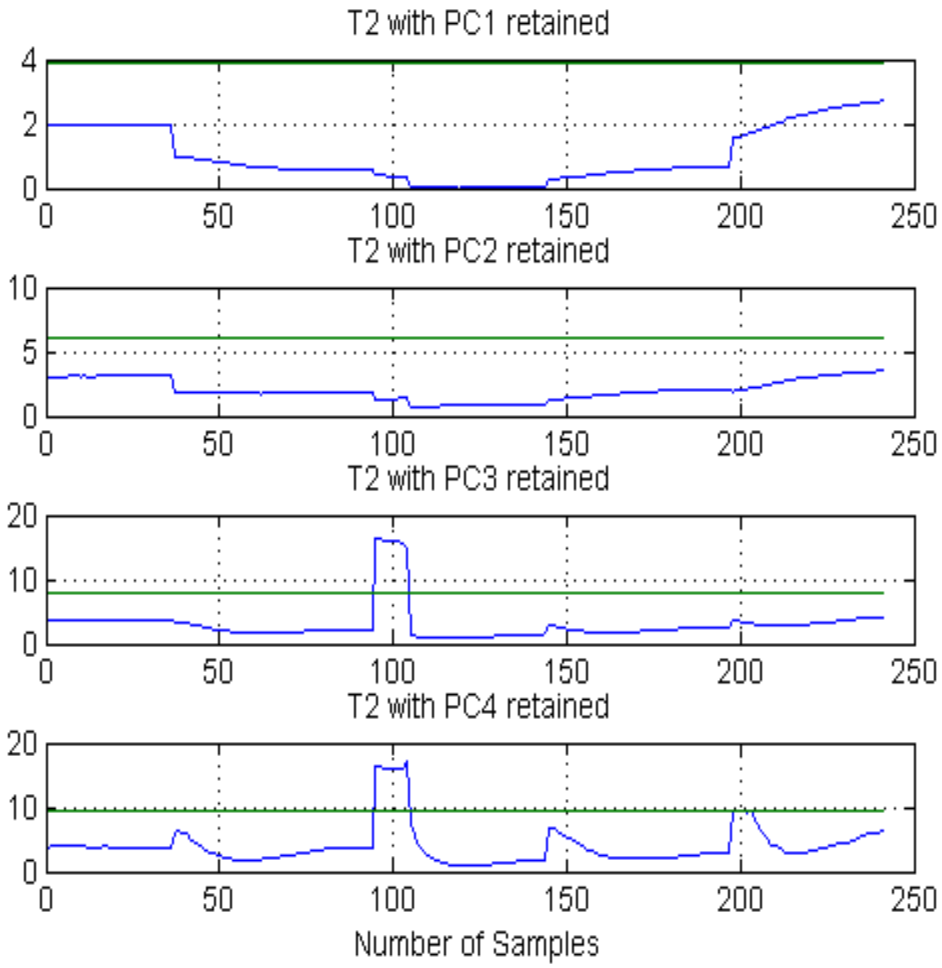

Figure 6.10: T^2 statistics plot for F3 of the thermal noisy data for various PCs retained.

The two plots show that fault 3 could be detected with the retention of only all (four) of the PCs as the rest of the combinations failed to detect fault 3.

The case was different with the noisy data, the retention of three PCs and four PCs could detect fault 3 and not two PCs or one PC. The introduction of more noise in the F3 data pushed the fault to the score space hence the detection by the retention of the 3PCs. For fault sensitivity one could retain four PCs to detect fault 3 for the data. Three PCs could be retained for the same fault 3 in the noisy data to serve the dimension reduction factor which is also important.

Conclusion

All the three methods used to determine the numbers of PCs are able to serve the purpose but fault SNR method works better than the two. This method was able to give the exact particular number of PCs to be retained which was sensitive to a particular fault which was not the case with the two Scree and the CPV methods. However the Scree and CPV methods are easy to apply for it only takes one performance to come up with the number of PCs to be retained. Fault SNR, all the PCs should be considered individually before getting the optimal number of PCs to be retained.

CHAPTER 7

RESULTS AND DISCUSSIONS

The complete laboratory thermal system is a multivariable system with two inputs and two outputs. Each input affects its respective output and the other, because each blower affects both sensors. Chapter indicated that the system is stable with bode plots and the root locus plots. The temperature change is affected by the environmental temperature hence the use of the results that were obtained on one particular day. All the three faults were artificially introduced on the thermal data affecting the same number of samples for detection purpose.

7.1 Thermal system Fault Detection

All the three faults introduced in the thermal system could be detected with different number of PCs retained using T^2 statistics. The introduction of noise on the system changed the results such that the fault that could only be detected by only retaining all PCs in case of F3, could now be detected by retaining three PCs as shown by Figures 5.8. This shows the fault shift, from the residual space to the score space, since T^2 statistics detects faults on the score space.

7.2 Comparison and Summary

The comparison of the three methods used to determine the number of PCs to be retained investigated in this dissertation has been given. Fault SNR is preferable over Scree test and CPV method because its ability to give the optimum number of PCs that is sensitive to a particular thermal system fault. This gives the superior performance of fault detection for different kinds of laboratory thermal system faults.

7.3 Conclusion and Recommendation

7.3.1 Fault Detection Using PCA

The fault detection method based on PCA shows that the number of PCs greatly affects the ability of fault detection. It was shown that the optimum number of PCs, which maximizes the sensitivity of the fault detection, depends on the kind of fault.

T^2 statistics was able to detect three different faults in the laboratory thermal system. Different retained number of PCs that were sensitive to different types of faults, some with the dimension reduced and the other with all the PCs (No dimension reduction).

It was also shown that all the introduced faults on the system were able to be picked up by the smaller number of PCs retained, but the insignificant faults, such as F3, which could occur due to environmental temperature and humidity change can only be detected by retaining all the PCs. If the reference data set of a faulty operation is available, optimization of the number of PCs is possible based on the fault SNR for process faults. This can be more useful if a certain type of fault repeatedly occurs in a plant.

7.3.2 Scree and Cumulative Percent Variance Methods

Figure 6.2 and figure 6.4 show that with T^2 Statistics technique, these two methods are not good to be used for determination of PCs for fault detection using PCA. Some faults cannot be detected if used hence not reliable.

7.3.3 Fault Signals to Noise Ratio (SNR) Method

Fault SNR was defined as an index of fault detection. The results from this method show that it has the capability to determine the number of PCs to be retained which are sensitive to a particular system fault. Amongst the three methods, fault SNR is the method that can be recommended when PCA is used for fault detection in a system with priori information, according to the results in this dissertation.

7.3.4 Future Work

Future work is needed to come up with a method of determining the number of PCs to be retained without having the priori information of the system faults. The method should be able to reduce the dimension at the same time be sensitive to a particular fault if not all types of faults that can occur in the system

APPENDIX A

MATLAB CODE FOR PCA

```
Function [signals, PC, V] = pca1 (data)
% PCA1: Perform PCA using covariance.
% data - MxN matrix of input data
% (M dimensions, N trials)
% signals - MxN matrix of projected data
% PC - each column is a PC
% V - Mx1 matrix of variances
[M, N] = size (data);
% subtract off the mean for each dimension
mn = mean (data, 2);
data = data - repmat (mn, 1, N);
% calculate the covariance matrix
covariance = 1 / (N-1) * data * data';
% find the eigenvectors and eigenvalues
[PC, V] = eig (covariance);
% extract diagonal of matrix as vector
V = diag (V);
% sort the variances in decreasing order
[junk, rindices] = sort (-1*V);
V = V (rindices);
PC = PC (: rindices);
% project the original data set
signals = PC' * data;
```

Computing PCA through SVD

```
function [signals, PC, V] = pca2 (data)
% PCA2: Perform PCA using SVD.
% data - MxN matrix of input data
% (M dimensions, N trials)
% signals - MxN matrix of projected data
% PC - each column is a PC
% V - Mx1 matrix of variances
[M, N] = size (data);
% subtract off the mean for each dimension
mn = mean (data, 2);
data = data - repmat (mn, 1, N);
% constructs the matrix Y
Y = data' / sqrt (N-1);
% SVD does it all
[u,S,PC] = svd(Y);
% calculate the variances
S = diag(S);
V = S * S;
% project the original data
signals = PC' * data;
```

APPENDIX B

SYSTEM IDENTIFICATION

This is a well established field with various numbers of methodologies which are divided into two main methods which are Parametric and non-Parametric methods [44, 45].

Parametric methods deal with relatively small number of system parameters, where these are optimized according to particular objectives. While non-parametric method are easy to use which makes them to be used where less structure is imposed on the model [44].

Some of the parametric methods are approaches found in the predictor family [46] and subspace approaches which are done in [47]. Non-parametric methods are statistical which include correlation and spectral analysis methods [48].

A number of statistical methods have not been implemented in the industry because of lack of probability information. These methods as complex as they are, they reduce to the same least square calculation as is the prediction error methods when the common probability assumptions are used. The Maximum Likelihood estimation and the Bayesian estimation methods are some of them.

B.1 Closed Loop System Identification Methods

Most industries as of late have been implementing the closed loop system identification and control related identification [49].

The assumptions that the systems controller is linear and the processes are SISO are most used by the users of the closed loop identification. This is not applicable to MPC applications because they are non-linear and most of the times MIMO. The MIMO and non-linear have been worked on [50, 44], where the use of closed loop system identification was found to have many advantages over open loop identification.

There are three ways that are used to classify the closed loop system identification [46].The direct approach, indirect approach and the joint input-output approach. Every closed-loop system identification is associated with the three, be it parametric or non-parametric methods.

The figure below shows the structure of the feedback system

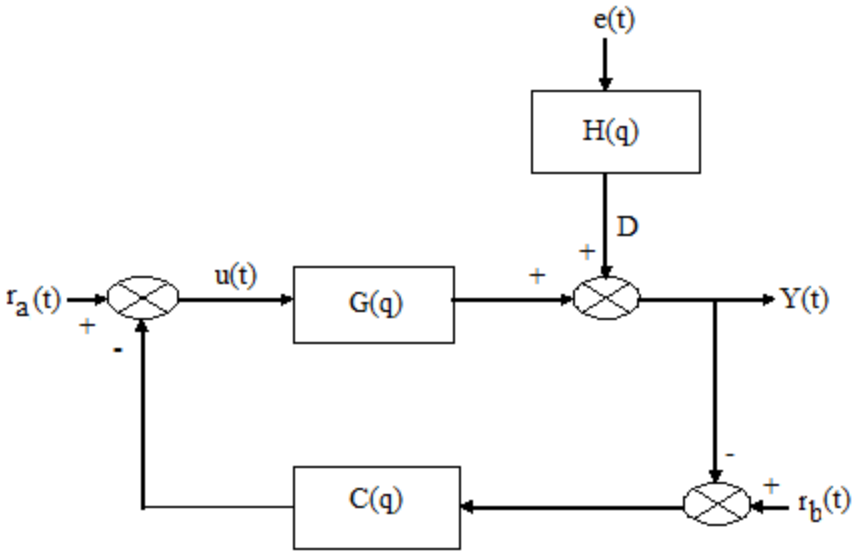

Figure B.1: The Structure of a Feedback System.

The equation describing the figure above could be written as:

$$y(t) = G(q^{-1}, \theta)u(t) + H_0(q^{-1}, \theta)e(t) \quad \text{B.1}$$

y (t) being the plant output, *u (t)* the plant input and *e (t)* the white noise while H_0 is the linear filter. H_0 is assumed to be stably invertible and monic such that $H_0(q) = \sum_{k=0}^{\infty} h(k)q^{-k}$, $h_0 = 1$ [52]. The white noise presented by *e (t)* in the figure, has a mean of zero and a covariance of P_e.

The design excitation signal r_a *(t)* is imposed on top of the process, the set point while the additional signal r_b *(t)* is designed for external excitation signal imposed on the controller. The symbol q^{-1} denotes the discrete time shift operator which can be expressed as $q^{-1}u(t) = u(t-1)$. The dynamic characteristics of the system model are denoted by θ which are realized by parametric coefficients.

The consistency of the approaches is concerned with the bias of parameter estimates while the efficiency is concerned with the asymptotic variance of the plant estimate [53].

B.1.1 Plant Estimates and Variance Expressions

Closed loop and open loop system, their variance expressions for plant estimates have been derived in [46, 54] for a plant estimate is defined as $\hat{G}$ (jw), and the covariance becomes:

$$\text{Cov}\,[\hat{G}\,(\text{jw})] \approx \frac{n\emptyset_D(w)}{N\emptyset_u(w)} \quad \text{B.2}$$

The equation above $\emptyset_D$(w) is the spectrum of disturbance, $\emptyset_u$(w) is the input spectrum, *n* the model order and *N* the number of samples. This equation indicates that the asymptotic variance of the plant estimate $\hat{G}$(jw) is directly proportional to the signal to noise ratio at any frequency. The equation is asymptotic in both *N* and *n*. For the closed loop plant estimate the variance expression was done in [54] and given as follows;

$$\text{Var}\,[\hat{G}(\text{jw})] \approx \frac{n\emptyset_D(w)}{N|S(w)|^2\,\emptyset_r\ (w)} = \frac{n\emptyset_D(w)}{N\emptyset_u(w)} \quad \text{B.3}$$

Given that $\emptyset_u^r$(w) = |s (w) $|^2\emptyset_r$(w) is the spectrum of the input signal coming from the external excitation that gives $u^r = s_0(q^{-1})r$ and

$$S_0(q^{-1}) = [1 + \text{G}\,(q^{-1})\,\text{C}\,(q^{-1})]^{-1} \quad \text{B.4}$$

This is the sensitivity function. Gevers et al [54] have shown that

Esmaili et al [55] made the conclusion using the above results for the open loop and closed loop systems that when the output power is limited, closed loop identification will generally give better identification (lower Var [$\hat{G}$(jw)] for the same output variance) than the open loop identification. This good part of closed loop identification would be found if one chooses the spectrum of the designed excitation signal, *r*, to be in a way that its contribution to the input *u* is the same as the input signal of the open loop situation as shown below;

$$\emptyset_u^r = |\text{S}|^2\emptyset_r = \emptyset_{u|open-loop} \quad \text{B.5}$$

B.1.2 Bias Distribution with Estimated System Parameters

The direct, indirect and joint input-output closed loop identification approaches consistency can be explained by analyzing estimated system parameter biases. Ljung [46] used the predictor error method for bias analysis to get the parameter estimates.

Direct Approach: In [45] it was shown that by applying the direct approach in closed loop identification, the parameter estimates could be given as;

$$\hat{\theta}_N \to \infty =$$
$$argmin_\theta \frac{1}{2\pi}\int_{-\pi}^{\pi}\{|G(jw) + B(jw,\theta) - G_\theta(q,\theta)|^2\,\emptyset_u(w) + |H_0(jw) - H(jw)|^2\emptyset_e^r(w)\}\,\frac{1}{|H(jw)|^2}\,\text{dw}$$

B.6

Where $B(jw,\theta)=|H_0(jw) - H(jw)|^2 \frac{\emptyset_u^e(w)}{\emptyset_u(w)}.\frac{\lambda_0}{\emptyset_u(w)}$ B.7

By inspection of the two equations above Forssel and Ljung [45] concluded that the bias$B(jw,\theta)$, which is realised by implementing the direct approach in closed loop systems. This can be small in the frequency ranges when the following conditions hold:

$H(jw)$, The estimated noise model is accurate, this is when $[\bar{\sigma}(H_0 - H)]$ is small.

$[{\bar{\sigma}(\emptyset_u^e)}/{\bar{\sigma}(\emptyset_u)}]$ which is the feedback contribution to the input spectrum is small.

The signal to noise ratio is high, this when $\frac{\lambda_0}{\emptyset_u(w)}$, is small.

Indirect approach: When the controller and some extra input or reference signals are known, then the indirect approach can be used for closed loop system identification [45]. This approach can also be used with the nonlinear feedback though most of the times linear feedback is used. The indirect approach for closed loop system can be given as:

$y(t) = G^c(q,\theta)\, r(t) + H(q)\, e(t)$ B.8

$G^c(q,\theta)=G(q,\theta)\,(1 + G(q,\theta)C(q))^{-1}$ B.9

One can identify the closed loop system and compute the estimates G_N (q, θ) of the open loop system if the controller C (q) is known and the reference signal/set point, *r (t)*, is measurable. This can be realized by solving the equation below [45]:

$\hat{G}_N^c = (1+ \hat{G}_N(q, \theta)C(q))^{-1}\hat{G}_N(q,\theta)$ B.10

The parameter estimate expression for the indirect approach was derived in [45] which is:

$$\hat{\theta}_N \to \infty = argmin_\theta \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \frac{G^c(jw) - G^c(iw,\theta)}{1+G_\theta(jw,\theta)C(jw)} \right| \frac{|S(jw)|^2 \emptyset_r(w)}{|H(jw)|^2} \, dw \quad \text{B.11}$$

In the equation above it can be concluded that the indirect approach in closed loop identification, and these conditions should be present [45].

It gives constant plant estimates if the parameterisations are flexible enough.

When not all the plant dynamics were identified, the plant estimate will minimise the mismatch between the nominal plants and also minimise the model sensitivity function.

Joint Input-Output Approach: With this approach, a model structure can be defined [45] as below:

$$\begin{bmatrix} y(t) \\ u(t) \end{bmatrix} = \begin{bmatrix} G^c(q,\theta) \\ S^i(q,\theta) \end{bmatrix} \text{r (t)} + \text{H (q)} \begin{bmatrix} e(t) \\ d(t) \end{bmatrix} \quad \text{B.12}$$

Where S^i presents the sensitivity functions of the input while *d (t*) and *e (t)* are independent noise sources.

In this approach, it is assumed that the input *u (t*) is generated by unknown controller, C_u(q), which gives the following form:

$$\text{u (t)} = \text{r (t)} - C_u(\text{q}) \text{ y (t)} \quad \text{B.13}$$

In [45] a parameter estimate expression was derived for the joint input-output approach with a model structure defined in equation A.12 as:

$$\hat{\theta}_N \to \infty = argmin_\theta \frac{1}{2\pi} \int_{-\pi}^{\pi} \{|G(jw) + B(jw) - G_\theta^c(q,\theta)|^2\} \frac{\emptyset_u(w)}{|H(jw)|^2} \, d \quad \text{B.14}$$

$$B\,(jw) = G\,(w)\, \emptyset_{uu}(w)\, \emptyset_u^{-1}(w) \quad \text{B.15}$$

For the parameter estimate in equation B.14 above that was derived for the joint input-output approach in [45] the following conditions should hold:

This approach will give consistent estimates of G (q) if the parameterisations of G^c and S^i are flexible.

The controller should not be known.

Regardless of the disturbance *e (t)*, the joint input-output approach gives consistent estimates of *G (q).*

B.2 System Identification Using Prediction Error Framework

This was proposed by Ljung [46], the identification formulations have been centered on PEM paradigm. The good thing with PEM is that convergence and asymptotic variance are well established [57], and the disadvantage of PEM is the complex parameterizations and non-convex optimization.

The application of PEM to closed loop was first proposed by ASYM method developed by Zhu [58].

The main aim behind the PEM is to estimate system parameters, so that the effect will minimize a prediction error objective function defined as follows:

$$\varepsilon(t,\theta) = y(t) - \hat{y}(t|\theta) \quad \text{B.16}$$

Data set Z^N *={y (1), u (1), y (N), u (N)}* B.17

These prediction errors can be computed for t = 1, 2... N. An accurate system model is a model that is accurate at predicting accurate outputs, thereby minimizing the prediction error [46]. There exist two main approaches that are being used in minimizing the prediction errors. There is one approach which is from a scalar valued norm or criterion function that measures the size of $\varepsilon(t,\theta)$ and then minimise the norm analytically. A well known criterion being used to minimize the prediction error norm is the least-squares criterion. Another approach to minimize the prediction error, is to demand that$\varepsilon\left(t,\hat{\theta}_N\right)$, be uncorrelated within a given data sequence. This requires that certain projections of $\varepsilon\left(t,\hat{\theta}_N\right)$ are zero [46]. In the following literature survey on PEM, the associated least-squares estimation methodology will be elaborated further, and the criteria involves the prediction error norm minimization will be discussed.

Linear model structures are good in describing basic linear and non-linear systems [46]. A predictor of a linear regression function, which is linear in the system parameter,θ can be described as;

$$\hat{y}(t|\theta) = \emptyset^T(t)\theta + \mu(t) \quad \text{B.18}$$

In which $\emptyset^T$, is the vector of regressors and $\mu(t)$ is a known dependent vector. From the prediction error equation 2.16, this gives;

$$\varepsilon(t,\theta) = \text{y (t) - } \emptyset^T \text{ (t)}\theta \qquad \text{B.19}$$

In which the known data dependent vector $\mu(t)$ is taken for notational simplicity. Assume a quadratic norm, $\frac{1}{2}\varepsilon^2\left(t,\hat{\theta}_N\right)$ to measure the prediction error. One can then define the linear regression equation B.18 [46] for the least squares criterion as:

$$\hat{\theta}_N^{LS} = \text{argmin } V_N(\theta, Z^N) = \left[\frac{1}{N}\sum_{t=1}^{N}\emptyset(t)\emptyset^T(t)\right]^{-1}\frac{1}{N}\sum_{t=1}^{N}\emptyset(t)y(t) \qquad \text{B.20}$$

Variants of the least squares estimation algorithm for parameter estimation in the PEM frameworks exist. Recursive Least-Squares (RLS) with forgetting factors [59], Robust RLS (RRLS) [60] and the weighted LS [44] are typically used for parameter estimation in the PEM framework.

B.2.1 Prediction Error Method

The prediction error method (PEM) deals with the estimation of linear system models by directly using the prediction error as a model performance and quality norm. In general the linear model used for system identification in the PEM framework can be defined as [44].

$$\text{A (q) y (k)} = \frac{B(q)}{F(q)}\text{u (k)} + \frac{C(q)}{D(q)}v(k) \qquad \text{B.21}$$

Most model structures are derived from the general linear structure in the above equation. The Autoregressive with Exogenous input (ARX) models and Autoregressive Moving Average with Exogenous input (ARMAX) models [46] are the most commonly used model structures that are concerned with the output feedback. The ARX model is the most widely applied linear dynamic model. Many real world processes structure match the ARX model. It easy to compute the parameters of ARX, this makes ARX to be popular. The parameters can be estimated by a linear least squares technique because the prediction error is linear in the parameters [44]. On the other hand parameter estimation of ARMAX is complicated. This makes the extended least squares algorithm to be necessary in solving the non linear optimization of the RMAX model parameters [44]. The disadvantages of the non linear optimization approach for solving model parameters are the existence of local optima and high computational demand. This is the implementation of the Recursive Least Squares Algorithms (RLS) [44].

Following the discussion on the variations of LS algorithms for estimating parameters in the PEM framework for different linear models used in the system identification and other parameter estimation methods.

B.2.2 RLS Method with a Static Forgetting Factor

A Recursive Least Squares (RLS) method using the direct approach to closed loop system identification was proposed by Eker [61]. The modeling of a three electromechanical system was done by using ARX model structure by Eker. The main advantages of RLS algorithms over LS are the fast parameter convergence and easy numerical solutions. The RLS method gives a consistent modeling accuracy over a wide range of operating conditions and is recommended as the best linear unbiased estimate [61]. The computational effort of the LS grows with the number of samples collected when it is run online in real time. This algorithm requires the constant computation time for each parameter update therefore this is perfectly suited for online parameter estimation [44].

The estimated ARX model output can be defined as follows;

$$\hat{y}(k) = \varphi^T(k)\,\hat{\theta}(k-1) \qquad \text{B.22}$$

The prediction error can also be defined as the following equation when the RLS uses the prediction error for model parameters update.

$$\varepsilon(k) = y(k) - \varphi^T(k)\,\hat{\theta}(k-1) \qquad \text{B.23}$$

And when updating system parameters the following equation is used:

$$\hat{\theta}^{RLS}(k) = \hat{\theta}(k-1) + P(k)\,\varepsilon(k)\varphi(k) \qquad \text{B.24}$$

In the equation 2.24, *P (K)* is the estimator covariance matrix which can be updated as follows:

$$P(k) = \frac{1}{\lambda} P(k-1)\left[I_p - \frac{\varphi(k)\varphi^T(k)(k-1)}{\lambda - \varphi(k)\varphi^T(k)(k-1)}\right] \qquad \text{B.25}$$

λ Is the static forgetting factor that is used in the RLS algorithm while the subscript p is the rank of the identity matrix. The forgetting factor determines the convergence speed; this makes the decreasing values of λ to give an increase in parameter convergence. But when the value of λ is made too small, it increases noise susceptibility. So [62] recommended that the range of λ should be $0.98 \leq \lambda \leq 0.995$. Eker [61] recommended that the initial values of *P (0)* and $\theta(0)$ should be chosen as follows:

P (0) = αI_Z where $0 \leq \alpha \leq 10^7$ and $\hat{\theta}(0) = 0$.

A fourth order ARX model was identified in [61] where the parameters were estimated via the RLS algorithm with a static forgetting factor.

B.3 Bi –loop Forgetting Factor RLS Method

This algorithm was proposed in [59], to overcome the weaknesses of the RLS static forgetting factor which has a slow tracking capability and high prediction errors. The large tracking errors at each sampling instant that leads to slow convergence of the RLS algorithm.

The bi-loop forgetting factor recursive least squares algorithm improves the tracking of time varying parameters as compared to the RLS implementation, with the same sampling rate. The BLFRLS is in the principle two forgetting factor recursive least squares (FRLS) algorithm where the outer loop algorithm computes parameter estimates at every sampling instant, while the inner loop (IFRLS) recursively recalculates and defines the parameter estimates in the N amount of time. This reduces the large tracking errors associated with the RLS algorithm. In essence IFRLS is exactly the same in principle as the RLS method by Eker. At sampling time $t = k$, the initial values of IFRLS can be given as:

$y_0 = y(k)$,

$x_0 = x(k)$

$\theta_0(0) = \theta(k)$ B.26

$P_{in}(0) = P(k-1)$

As the initial values are shown in the equations above, the estimator covariance matrix P_{in} and the parameter estimates $\hat{\theta}_{in}$ are iteratively calculated just as for RLS algorithm.

The closed loop parameter estimation using the PEM framework was proposed by YU and Shih in [59]. The efficiency of the proposed algorithm was simulated for three scenarios while using the ARX model structure. The three are estimation of slow varying dynamic, the tracking of a sinusoidal parameter and the estimation of the fast varying parameter. [59] Also shows that when the simulation settings are N = 10 and λ = 0.98, the BLFRLS algorithm can handle abrupt parameter changes more efficiently and effectively than the RLS algorithm.

B.4 Robust Recursive Least Squares Method

This method was first done by Chao et al [60]. The main idea was to use it for online estimation of time varying parameters of an AR model using a weighted LS method with forgetting factors. The sum of the squared prediction errors is minimized when using the conversional LS estimation, where the distribution of the prediction errors is considered to be Gaussian in the LS procedure. Since the automatically obtained data inevitably carry the some false data and gross errors which may result in a different prediction error distribution, with the remote sensing. This may lead to the deteoration of the efficiency of the LS procedure because the LS procedure weights all prediction errors equally. The robust solution in [60] assigns a weight as a function to the prediction errors; in this case a loss function is used to assign more weight to the bulk of small prediction errors and less weight to the gross errors which are called outliers. The RRLS is different from the conventional RLS algorithm in that it inserts a nonlinear transformation function for the prediction errors. When the large outliers are transformed and a small weight is assigned to these outliers, this dramatically reduces the bias of the RLS estimation [60].

The case is different with conventional LS estimation in which its objective is to minimize a cost function which is:

$$J(\hat{\theta}) = \frac{1}{2}\mathrm{E}[\varepsilon^2(t,\theta(t))] \qquad \text{B.27}$$

The reformulation of the cost function changes with the nonlinear transformation of the outlier to ensure robust estimation, and it becomes:

$$J(\hat{\theta}) = \frac{1}{2}\mathrm{E}\left[\rho\{\varepsilon^2(t,\theta(t))\}\right] \qquad \text{B.28}$$

Where ρ (.) is a nonlinear loss function which suppresses the out-liers. In [60] it is stated that the loss-function be like the quadratic function for small prediction error values of argument. It also states that it's a requirement that the derivative of ρ(.) should be bounded and continuous, i.e $\psi = \rho'$ (.). When it is bounded then no single observation can have an arbitrary large influence on the parameter estimation. Its continuity makes that rounding and quantization errors to have no major effect.

In the same [60] the loss function, the derivative loss function and the weighing factor respectively are as below:

$$\rho(\varepsilon_i) = \begin{cases} \frac{\varepsilon_i}{2} & |\varepsilon_i| \le k_1\sigma \\ k_1.|\varepsilon_i| & k_1 < |\varepsilon_i| \le k_2\sigma \\ k_2 & |\varepsilon_i| > k_2\sigma \end{cases} \qquad \text{B.29}$$

$$\Psi(\varepsilon_i) = \begin{Bmatrix} \varepsilon_i & |\varepsilon_i| \le k_1\sigma \\ k_1.sgn(.) & k_1 < |\varepsilon_i| \le k_2\sigma \\ 0 & |\varepsilon_i| > k_2\sigma \end{Bmatrix} \quad \text{B.30}$$

$$\omega(\varepsilon_i) = \begin{Bmatrix} 1 & |\varepsilon_i| \le k_1\sigma \\ \frac{k_1}{\varepsilon_i} & k_1 < |\varepsilon_i| \le k_2\sigma \\ 0 & |\varepsilon_i| > k_2\sigma \end{Bmatrix} \quad \text{B.31}$$

In all the three equations above k_1 and k_2 are nonlinear tuning constants. The good values that are reasonable and recommended to be used are k_1 = 1.5 and k_2 = 2.5[63].

The effect of the weighting factor is to assign less weight to outliers which can improve the accuracy of the parameter estimates. This function also depends on the parameter variance , σ^2, where σ depends on θ which is determined by ω. The same function is also determined by iteration where $\omega = 1$ which is the same as that of RLS method [60].

The inclusion of the dynamic forgetting weighting factor is proposed in [60] for weighting more recent data more heavily in the computation of parameter estimates. The dynamic factor is:

$$\lambda = \beta^{t-1} \quad \text{B.32}$$

$$\hat{\theta}^{RRLS}(t+1) = \hat{\theta}^{RRLS}(t) + \omega(t+1)P_t X_{t+1}[\lambda + \omega(t+1)P_t X_{t+1}X_{t+1}^T]^{-1} \text{ x } [y(t+1) - X_{t+1}^T \hat{\theta}^{RRLS}(t)] \quad \text{B.33}$$

And P which is the estimator covariance matrix becomes:

$$P_{t+1} = \lambda\{I - \omega(t+1)P_t X_{t+1} \text{ x } [\lambda + \omega(t+1)P_t X_{t+1}X_{t+1}^T]^{-1} X_{t+1}^T\}P_t \quad \text{B.34}$$

The implementation of the RRLS algorithm on a real data in flood forecasting was done in [60]. And it was concluded that the RRLS algorithm produces less biased estimates than the conventional RLS algorithm, this was when AR model were estimated. It was also proved that RRLS model is more robust to outliers in real time.

B.5 Kalman Filtering Method

Kalman filter is also used to estimate the model parameters instead of the conventional state estimation. This was explained in [64] where closed loop system identification was done using the direct approach in the PEM framework. This filtering method is almost the same as the RLS approach. The Kalman filtering

is mostly applied as an observer for the estimation of states not parameters. In the state space formulation, the parameter estimation can be presented as:

$$\theta(k+1) = \theta(k) + v(k),$$

$$y(k) = x^T \theta(k) + e(k), \qquad \text{B.35}$$

Where $v(k)$ is a n-dimensional vector representing white noise with an *n x n* dimensional covariance matrix V, where *n* is the number of parameters. The time variance of the parameters are modeled as a random walk drift [65]. The covariance matrix is usually chosen to be diagonal. The strength of the time variance of the *n* individual parameters is the diagonal entries. If the parameters are known, the corresponding entry for V should be large [44].The fact that each parameter has its own forgetting factor makes the Kalman filtering method better than the RLS algorithm [64]. This is an advantage because it makes it possible to control the convergence of individual parameters, the statistical characteristics of the parameters bounds the convergence of each parameter, this is done by setting the forgetting factor accordingly. The V can be set to ξI if no statistical knowledge is available about the individual parameters. The forgetting factor λ = 1 is equivalent to V = 0.The parameter estimation by the kalman filter algorithm can be formulated as below:

$$\hat{\theta}(k) = \hat{\theta}(k-1) + \gamma(k)e(k)$$

$$e(k) = y(k) - x^T(k)\hat{\theta}(k-1)$$

$$\gamma(k) = \frac{1}{x^T(k)P(k-1)x(k) + {}^{1}/_{q(k)}} P(k-1)x(k) \qquad \text{B.36}$$

$$P(k) = (I - \gamma(k)x^T(k))Pk - 1) + V$$

Where $\gamma(k)$the adaptation is factor, and in this case it is a Kalman gain. The P matrix does not increase exponentially in the Kalman filter algorithm as it is with the RLS algorithm [64]. Linearly $P(k) = P(k-1) + V$ when there is non-persistent excitation. When the signal of excitation is able to excite the process dynamics of the plant over the entire frequency band operation that is when the persistent excitation is realized.

Identification of the AR model of a nonlinear aerospace launch vehicle system was done when testing the method [64]. In conclusion it was found that the proposed method of the estimating system parameters by the implementation of a Kalman filter increased the precision and convergence of the model parameters in comparison to the conventional RLS algorithm.

A technique to identify system parameters of nonlinear dynamic systems which is well known is the use of extended Kalman filter. In this technique, parameters are treated as states, which is the same as it was discussed in [64]. Although the extended Kalman filter approach has the restriction in that the nonlinear state equation needs to be differentiated with respect to each state variable [66]. There is also a problem with extended Kalman filter (EKF) that makes it difficult to implement, it is difficult to tune and is only reliable for systems that are almost linear on the time scale of the updated intervals [67].The unscented Kalman Filter (UKF) was proposed by Julier et al [67] to help in the solving of the problem of nonlinear non-differentiable process model. This is a linear estimator equivalent to the Kalman filter for linear systems in its performance, but the UKF can be generalized to nonlinear systems without going through the steps necessary in the EKF approach [67]. UKF does not approximate nonlinear functions as the EKF but uses a set of deterministically chosen weighted sample points to get the estimator of the state variables [66]. This makes it UKF a method that can be applied in nonlinear system with discontinuities if used for estimation. Parametric discontinuities can happen in the context of an industrial process, unplanned plant shutdowns or maintenance. This may result in the loss of accurate parameter tracking of the conventional Kalman filter and EKF estimation approaches.

The UKF method was also implemented by Araki et al [66] for estimating unknown parameters of a 2-link under actuated acrobat robot. In this case the unknown parameters were taken as unknown states for estimation. The UKF was developed based on the fact that it is easier to approximate a probability distribution than to approximate an arbitrary nonlinear function or transformation. It follows that a nonlinear transformation function is applied to a set of points for which the sample mean and sample covariance are $\hat{x}(k|k-1)$ and$P(k|k-1)$. A new set of predicted mean and covariance values result from the transformation of this set. Though this method performs like the Monte Carlo method, the samples are not randomly chosen, but still the specific information is captured about the distribution of the states. One can get the minimum mean square error (MMSE) state for estimator for a nonlinear system as:

$$x\,(k+1) = f[x(k), u(k), k],$$

$$z(k+1) = h[x(k+1), u(k+1), k+1] + \omega(k+1) \qquad \text{B.37}$$

[70] $x(k)$ Is the state of the system at time step$k$, $u(k)$ is the input vector, $z(k)$ is the observation vector while $\omega(k)$ is the additive measurement noise. $f(.)$ and $h(.)$ are the nonlinear system and measurement functions respectively. It is assumed that the process noise and the measurement noise have zero mean [70] and the covariances are given as:

$$E[v(k)v^T(j)] = \delta_{kj}Q(k),$$

$$E[\omega(k)\omega^T(j)] = \delta_{kj}R(k), \forall k, j \quad \text{B.38}$$

$$E[v(k)\omega^T(j)] = 0.$$

The UKF use $2n + 1$ regression points,x_i in state space with weights $W_i(i = 1, \dots, n)$ as:

$$x_0(k|j) = \hat{x}(k|j)$$

$$W_0 = \frac{2k}{2(n+\hbar)}$$

$$x_i(k|j) = \hat{x}(k|j) + \left(\sqrt{(n+\hbar)P(k|j)}\right)_i$$

$$W_i = \frac{1}{2(n+\hbar)}.$$

$$x_{i+n}(k|j) = \hat{x}(k|j) - \left(\sqrt{(n+\hbar)P(k|j)}\right)_i \quad \text{B.39}$$

$W_{i+n} = \frac{1}{2(n+\hbar)}$ Of the transformed samples:

In the equation B.39 $\left(\sqrt{(n+k)P(k|j)}\right)_i$ is the i-th row or column of $\left(\sqrt{(n+\hbar)P(k|j)}\right)$. $\hbar$ is used to refine the higher order moments of the approximation as an extra degree of freedom, with $\neq 0$, [69] when choosing the regression pointsx_i. The following is the prediction procedure that is followed when given the set of samples that are generated by B.39;

Every $\mathcal{X}$ is passed through the process model to give a set of transformed samples:

$$\mathcal{X}_i(k+1|k) = f[\mathcal{X}_i(k|k-1), u(k), k] \quad \text{B.40}$$

The computed prediction mean is:

$$\hat{x}(k+1|k) = \sum_{i=0}^{2n} w_i\, x_i\,(k+1|k). \quad \text{B.41}$$

The computed covariance becomes:

$$P(k+1|k) = \sum_{i=0}^{2n} w\{x_i(k+1|k) - \hat{x}(k+1|k)\}x\{x_i(k+1|k) - \hat{x}(k+1|k)\}^T \quad \text{B.42}$$

The calculation of the mean and the covariance is done by using the standard vector and matrix operations, this shows that the algorithm can be applied to any process model [67].

Discussion

Various system identification methods in particular in parameter estimation have been outlined in this section. The conventional LS method used for solving the parameter estimation under PEM framework is good but not good enough for online parameter estimation.

The online parameter estimation can be solved by using the RLS algorithm because the computational work of the LS method grows with the number of samples collected [61]. RLS method is suitable for the online use of in real time applications because it requires a constant time for each parameter update.

A bi-loop RLS method is an extension on RLS algorithm to address the problem of the trade off between the parameter convergence speed and the noise susceptibility of the RLS method [59]. The parameter estimation together with the forgetting factor is used to reduce the weight of influence of past sampled data in the parameter estimation process. These can also increase the convergence of the parameters to their true values.

RRLS method, the data outliers in this method does not affect the parameter estimation since it is robust [63]. To prevent the influence of the outliers, the prediction errors used with the parameter estimation process are transformed nonlinearly. The RLS method is similar to the Kalman filter approach that is used for state estimation.

EKF approach is used for estimating states and at the same time estimate the process parameters of a nonlinear system. The UKF was proposed to help in the weaknesses of EKF approach. UKF does not approximate nonlinear functions like the EKF but it uses a set of deterministically chosen weighted sample points to come up with the state variable estimator [66].

B.6 Identification by subspace

The subspace methods are relatively new methods of system identification whose origins are between the fields of numerical linear algebra, geometry and system theory [68]. Linear subspace identification methods deal with systems and models that can be represented in state space. MIMO systems are the ones which are suitable for state space models. It is difficult to find a numerically robust and canonical realization when applying the predictor error methods to state space models [44]. This is due to the alternative which is the full parameterizations of the state space model can involve a huge number of parameters. We can define the state space models for linear subspace identification as:

$$\begin{bmatrix} x_{k+1} \\ y_k \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} x_k \\ u_k \end{bmatrix} + \begin{bmatrix} K\omega_k \\ v_k \end{bmatrix} \quad \text{B.43}$$

With

$$E\left[\begin{pmatrix} \omega_k(t) \\ v_k(t) \end{pmatrix} \begin{pmatrix} \omega_k^T(\tau) & v_k^T(\tau) \end{pmatrix}\right] = \begin{pmatrix} Q & S \\ S^T & R \end{pmatrix} \delta_{t-\tau} \geq 0 \quad \text{B.44}$$

In the equations 2.43 and 2.44, the vectors $u_k \in \boldsymbol{R}^{m\text{x}1}$ and $y_k \in \boldsymbol{R}^{l\text{x}1}$

Are measurements taken at time instants of k for m inputs and l outputs of the process. The state space vectors of a system $\boldsymbol{x}_k$ and the unobserved vector signals $v_k \in \boldsymbol{R}^{l\text{x}1}$ and$\omega_k \in \boldsymbol{R}^{n\text{x}1}$. The vector v_k is the measurement noise and the vector ω_k is the process noise. **K** is the matrix for the Kalman gain while $\boldsymbol{A} \in \boldsymbol{R}^{n\text{x}n}$ is the system matrix and $\boldsymbol{B} \in \boldsymbol{R}^{n\text{x}m}$ is the input matrix. $\boldsymbol{C} \in \boldsymbol{R}^{l\text{x}n}$ Is the output matrix while $\boldsymbol{D} \in \boldsymbol{R}^{l\text{x}m}$ is a direct feed through matrix. The matrices $\in \boldsymbol{R}^{n\text{x}n}$, $\boldsymbol{S} \in \boldsymbol{R}^{n\text{x}l}$ and $R \in \boldsymbol{R}^{l\text{x}l}$ are the covariance matrices of the noise sequences are v_k and ω_k [1]. δ Is the kronecker delta while E is the expected operator.

The difficult thing about the subspace system identification is to come up with the order n of the system which is not known, the system matrices which are $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \boldsymbol{D}$ to a similarity transformation and estimate the covariance matrices which are$\boldsymbol{Q}, \boldsymbol{S}, \boldsymbol{R}$, of the measurement and process noise. The input u_k which is a large number of measurements and outputy_k, which are generated by the system which is unknown are required. There is a general assumption that the amount of data points go to infinity [68].

To identify a system using the subspace, two basic steps should be followed, the projection of certain subspaces generated from the measured data sets to find an estimate of the extended observability matrix $\hat{\Gamma}$ and an estimate of the states, $\widehat{\boldsymbol{X}}_k$**,** of the unknown system as the first step. While the second step involves the retrieving of the system matrices from the obtained extended observability matrix or estimated system states [68]. The following figure gives the details of the subspace algorithm two basic steps.

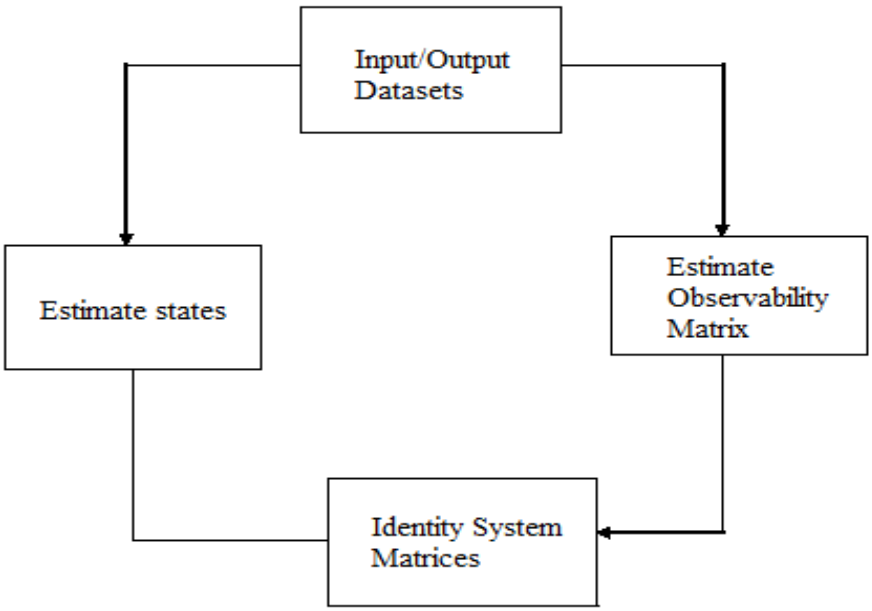

Figure B.3: steps to subspace Identification methods

The following input-output matrix equation which is the extended state space model of the equation B.43. It was used in the development of subspace system identification:

$$Y_f = \hat{\Gamma}_i X_i + H_i^d U_f + H_i^s M_f + N_f \qquad \text{B.45}$$

The term $\hat{\Gamma}_i$ is the extended observability matrix, where H_i^d and H_i^s are the deterministic lower block triangular Toeplitz matrix and the stochastic lower block triangular Toeplitz matrix. M_f And N_f are defined as the future block Hankel matrices formed with process noise ω_k and $v(k)$. the subspace methods states that the Hankel matrix containing the future outputs is related in the linear way to the Hankel matrix containing the future inputs and the future state sequence. The main idea of subspace identification is to recover the $\hat{\Gamma}_i X_i$ term of the equation 2.45. The knowledge of this term helps to come up with the system parameters, where the singular value decomposition of this term gives the system order. This is so because $\hat{\Gamma}_i X_i$ is rank efficient.

The projection of the row space of Y_f into the orthogonal complement of the row space of U_f is the first step in obtaining the estimate of the term $\hat{\Gamma}_i X_i$ as shown below:

$$Y_f (U_f^{\perp})^{-1} = \hat{\Gamma}_i X_i (U_f^{\perp})^{-1} + H_i^d U_f (U_f^{\perp})^{-1} + H_i^s M_f (U_f^{\perp})^{-1} + N_f (U_f^{\perp})^{-1} \qquad \text{B.46}$$

Where it follows;

$$Y_f(U_f^{\perp})^{-1} = \hat{\Gamma}_i X_i(U_f^{\perp})^{-1} + H_i^s M_f + N_f \quad \text{B.47}$$

Weighting the above equation with W_1 and W_2 ;

$$W_1[Y_f(U_f^{\perp})^{-1}]W_2 = W_1[\hat{\Gamma}_i X_i(U_f^{\perp})^{-1}]W_2 + W_1[H_i^s M_f + N_f]W_2. \quad \text{B.48}$$

The weight matrices together with U_f must be chosen in a way that the following conditions are met:

$$\begin{aligned} &rank\,(W_1\Gamma_i) = rank(\Gamma_i) \\ &rank\,[Y_f\big(U_f^{\perp}\big)^{-1}W_1\big] = rank(X_i) \\ &W_1\big[\,H_i^s M_f + N_f\big]W_2 = 0 \end{aligned} \quad \text{B.49}$$

The first two conditions guarantee that the system rank n, is kept after projection, while the third condition expresses the necessity of the weight W_2 to be uncorrelated with the noise sequences ω_k and $v(k)$[68]. From these conditions the weighted input-output matrix can be defined as:

$$O_i \overset{\text{def}}{=} W_1[Y_f(U_f^{\perp})^{-1}]W_2 \quad \text{B.50}$$

After singular value decomposition equation B.50 becomes;

$$svd(O_i) = (U_1 \quad U_2)\begin{pmatrix} S_1 & 0 \\ 0 & S_2 \end{pmatrix}\begin{pmatrix} V_1^T \\ V_2^T \end{pmatrix} \quad \text{B.51}$$

Where U_1 and U_2 are the unit matrices of output singular vectors , V_1 and V_2 are the unit matrices for input singular vectors while S_1 and S_2 are the eignevalues. From the equation 2.51 the following results are found:

$$n = rank(O_i) \quad \text{B.52}$$

$$W_1\Gamma_i = U_1 S_1^{1/2} \quad \text{B.53}$$

$$X_i(U_f^{\perp})^{-1}W_2 = S_1^{1/2}V_2^T \quad \text{B.54}$$

From the two equations above, it is possible to obtain the system parameter estimates by either using the extended observability matrix or the estimates of the system states. The system parameters A, B, C, D can be obtained by using the estimated system states and thereby solving the equation below using the least squares method.

$$\begin{pmatrix} \hat{X}_{i+1} \\ Y_i \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} \hat{X}_i \\ U_i \end{pmatrix} + \begin{pmatrix} \rho_w \\ \rho_v \end{pmatrix} \quad \text{B.55}$$

Where ρ_w and ρ_v are the residual matrices of the estimation process.

The least squares expression to be solved can be defined as in [68]:

$$min_{A,B,C,D} \sqrt{tr(E_{LS}^H E)} \quad \text{B.56}$$

Where $\begin{pmatrix} \hat{X}_{i+1} \\ Y_i \end{pmatrix} - \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} \hat{X}_i \\ U_i \end{pmatrix}$ and the trace tr is the sum of the diagonal elements while E_{LS}^H is the complex conjugate transpose ofE_{LS}.

The covariance matrices $Q, S, R,$ are estimated by using the residuals as follows:

$$\begin{pmatrix} Q & S \\ S^T & R \end{pmatrix}_i = \frac{1}{i}\left[\begin{pmatrix} \rho_w \\ \rho_v \end{pmatrix} (\rho_w{}^T \quad \rho_v{}^T)\right] \geq 0 \quad \text{B.57}$$

In this case i is a finite bias which disappears as it goes to infinity.

B.6.1 Subspace Methods

Subspace system identification methods are regarded as an alternative to PEM identification methods. It can be applied in MIMO without the need for special parameterizations that require significant prior knowledge and non-convex optimization [52]. When there is no unbiased estimate property of the subspace most subspace methods can not work, especially when used with closed loop system data, and with large data sets [52, 69]. The table below shows the main subspace methods that are used, W_1 and W_2 are the suitable weighting matrices that are chosen, for all subspace algorithms for LTI systems [68].

METHOD	$\boldsymbol{W_1}$	$\boldsymbol{W_2}$
N4SID	I_{li}	$(W_p(U_f^{\perp})^{-1})^{\dagger}W_p$
CVA	$[Y_f(U_f^{\perp})^{-1}][\ (Y_f(U_f^{\perp})^{-1})^{-1/2}]$	$[W_p(U_f^{\perp})^{-1})]^{\dagger}[(W_p(U_f^{\perp})^{-1})]$
MOESP	I_{li}	$[W_p(U_f^{\perp})^{-1})]^{\dagger}[(W_p(U_f^{\perp})^{-1})]$
Basic-4SID	I_{li}	I_j
IV-4SID	I_{li}	ϕ

Table B.1: Various subspace algorithms in a unifying frame work.

† denotes the Moore-Penrose pseudo-inverse. The state estimates $\widehat{\boldsymbol{X}}_{\boldsymbol{i}}$ are used in the first two algorithms to come up with the system matrices while the last three used the extended observability matrix $\hat{\Gamma}_{\boldsymbol{i}}$ which has the instrumental variables.

The five subspace system identification methods work well in the open loop identification [70 and 71] has more information on these methods. The N4SID was modified to work in closed loop system identification by Overschee and De Boor. The new subspace identification methods that are used in the closed loop systems will be investigated in this section.

B.6.1.1 Free Model Reduction Algorithm

Mathieu and Mohammed [72] proposed a subspace identification method for closed loop subspace system identification algorithm which can be divided into three steps, which are:

The first step consists of determining the Markov parameters of the control systems sensitivity, G_{cl}, which can be defined as follows:

$$G_{cl} = \begin{pmatrix} S_{r1u} & S_{r2u} \\ S_{r1y} & S_{r2y} \end{pmatrix}$$

$$= \begin{pmatrix} (I_m + CG)^{-1} & (I_m + CG)^{-1}C \\ G(I_m + CG)^{-1} & G(I_m + CG)^{-1}C \end{pmatrix} \quad \text{B.58}$$

The functions of sensitivity are defined for the closed loop system between the reference signals r_1, r_2 and the inputs and outputs u, y for the controller $C$and the plant G.

The impulse response of the system is identified from the results in step1 and this marks the beginning of step2. In this second step the finite amount of Markov parameters of the system G are estimated. The three equations below illustrate all that is involved in step 2.

$$\hat{\Gamma}_{cl/i} = L_i^T \begin{Bmatrix} -W_i^{-1}H_{c/i}\ \hat{\Gamma}_{G/i} & W_i^{-1}\ \hat{\Gamma}_{c/i} \\ (I_{ip} - H_{G/i}W_i^{-1}H_{C/i})\ \hat{\Gamma}_{c/i} & H_{G/i}W_i^{-1}H_{C/i}\hat{\Gamma}_{c/i} \end{Bmatrix} T^{-1} \quad \text{B.59}$$

$$\Delta_{cl/i} = T \begin{Bmatrix} \Delta_{G/i}W_i^{-1} & \Delta_{G/i}W_i^{-1}H_{c/i} \\ -\Delta_{C/i}H_{G/i}W_i^{-1} & \Delta_{\frac{c}{i}}(I_{ip} - H_{G/i}W_i^{-1}H_{C/i}) \end{Bmatrix} L_i \quad \text{B.60}$$

$$H_{cl/i} = L_i^T \begin{Bmatrix} W_i^{-1} & W_i^{-1}H_{C/i} \\ H_{G/i}W_i^{-1} & H_{G/i}W_i^{-1}H_{C/i} \end{Bmatrix} L_i \quad \text{B.61}$$

In the equations above $\hat{\Gamma}$ and Δ are the extended observability matrix and the reverse extended controllability matrix. In [72] the useful$\hat{\Gamma}_{cl/i}$, $\Delta_{cl/i}$ and $H_{cl/i}$ were derived as shown in the three equations above.

The matrices $W_i = (I_{im}H_{G/i}H_{C/i})$ and T are non singular transformation matrices and matrix L_i as well.

$$\begin{pmatrix} \hat{\Gamma}\Delta_{G/i} & 0 \\ 0 & \hat{\Gamma}\Delta_{C/i} \end{pmatrix} = \begin{pmatrix} -H_{G/i} & I_{ip} \\ I_{im} & H_{C/i} \end{pmatrix} \left(L_i \hat{\Gamma} \Delta_{cl/i} L_i^T \right) \begin{pmatrix} I_{ip} & -H_{C/i} \\ H_{G/i} & I_{im} \end{pmatrix} \quad \text{B.62}$$

One can get the term $\hat{\Gamma}\Delta_{G/i}$ from $H_{G/2i}$ as indicated in [77] and the equation below has it defined:

$$\hat{\Gamma}\Delta_{G/i} = (-H_{G/i} \quad I_{ip}) \begin{pmatrix} H_{21} & H_{23} \\ H_{41} & H_{43} \end{pmatrix} \begin{pmatrix} I_{ip} \\ H_{G/i} \end{pmatrix} \quad \text{B.63}$$

The third step which is the last one deals with determining the system order, together with a state space realization system. The order is determined using singular value decomposition method [73 and 47]. The proposed closed loop subspace identification method is used to identify a plant setup of two circular plates rotated by an electrical servo motor with flexible shafts [72]. This method was compared with the direct identification under PEM algorithm, initialized by N4SID estimate. The subspace method performed very well compared to the direct identification approach.

B.6.1.2 Parsim – E Algorithm

Subspace identification algorithms use a non-parsimonious model formulation with extra terms in the model that appear to be non-casual. These terms are conveniently included to perform subspace projection, but are the cause for inflated variance in the estimates, and partially responsible for the loss of closed loop causal terms are removed, making the model parsimonious. This method removes the condition of no correlation between the future and input u_k and the past innovatione_k, which is the case

for closed-loop data. The main idea of this method is accomplished by partitioning the extended state-space model in a row wise to exclude non-causal terms in the model. Below is a definition of the partitioned extended state-space model.

$$Y_{fi} = \Gamma_{fi} X_k + H_{fi} U_i + G_{fi} E_i \quad \forall i=1,2,\ldots,f \tag{B.64}$$

In this equation f denotes the future horizon.

The above equation may result into parsimonious model representation because it is very likely it is to be causal [74]. To eliminate $e(t) = \begin{pmatrix} w(t) \\ v(t) \end{pmatrix}$ in the innovation model through iteration, it is possible to reformulate the partitioned the extended state-space model as below:

$$Y_{fi} = \Gamma_{fi} L_z X_p + \Gamma_{fi} A_K^p X_{k-p} H_{fi} U_i + G_{fi} E_i, \quad \forall i = 1,2,\ldots,f. \tag{B.65}$$

Where p denotes the past horizon and

$$L_z \stackrel{\text{def}}{=} \left[\Delta_p(A_k, K)\Delta_p(A_k, B_k)\right]$$

$$\Delta_p(A, B) \stackrel{\text{def}}{=} [A^{p-1}B \ldots AB\ B] \tag{B.66}$$

$$A_k \stackrel{\text{def}}{=} A - KC \tag{B.67}$$

$$B_k \stackrel{\text{def}}{=} B - KD$$

$$Z_p \stackrel{\text{def}}{=} \begin{bmatrix} Y_P^T & U_p^T \end{bmatrix}$$

As p tends to infinity the second term on the right of the equation above tends to zero. The least squares estimates for the parsimonious can be given as below:

$$\begin{bmatrix} \hat{\Gamma}_{fi} L_z & \widehat{H}_{fi} \end{bmatrix} = Y_{fi} \begin{bmatrix} Z_P \\ U_i \end{bmatrix}; \ \forall i = 1,2,\ldots,f \tag{B.68}$$

These estimates are biased for closed loop identification according to Qin and Ljung [79]. They treated the estimated innovation as the known data, which in subsequent projections results in not requiring future inputs to be correlated with the past innovations. It was also derived that the least square estimate that does not require the future input u_k to be uncorrelated with the past innovation as e_k as below:

$$\begin{bmatrix}\hat{\Gamma}_{fi}L_z & \hat{H}_{fi} & \hat{G}_{fi}\end{bmatrix} = Y_{fi}\begin{bmatrix}Z_P \\ U_i \\ \hat{E}_{i-1}\end{bmatrix} \quad \text{B.69}$$

$$\hat{G}_{fi} = \begin{bmatrix}CA^{i-2}K & CA^{i-3}K \dots CK\end{bmatrix}$$

One can recursively calculate the innovation data as below:

$$\hat{E}_i = \begin{bmatrix}\hat{E}_{i-1} \\ \hat{E}_{fi}\end{bmatrix} \quad \text{B.70}$$

The comparative simulation studies were done in [79] between the $PARSIM-E$ and $MOESP$ algorithms. These were done on both open loop and closed data. The conclusion was that both data performs the same with no differences for open loop and very different in closed loop system identification. The best estimates were without bias were given by $PARSIM-E$ algorithm while $MOESP$ fails in closed loop identification.

B.6.1.3 SSARX ALGORITHM

This is one of the methods of subspace system identification algorithms, which is a Stacked outputs ARX (SSARX). This method makes use of CCA subspace identification theory and ARX modeling, it is able to use data in the open or closed loop.

Using this method the state space equations can be reformulated as below:

$$x(t+1) = \tilde{A}x(t) + \tilde{B}u(t) + Ky(t) \quad \text{B.71}$$

$$y(t) = Cx(t) + Du(t) + e(t)$$

Where the matrices $\tilde{A}$ and $\tilde{B}$ are;

$$\tilde{A} = (A - KC)$$

$$\tilde{B} = (B - KD) \quad \text{B.72}$$

Then the extended state space equation in this method can be reformulated as:

$$y_f(t) = \hat{\Gamma} x(t) + \hat{\Phi} u_f(t) + \hat{\Psi} y_f(t) + e_f(t) \quad \text{B.73}$$

The equation above is regarded as the stacked outputs of the ARX model which mostly have an infinite order.

In the equation the subscript f stands for the future horizon. And the matrices $\hat{\Gamma}$, $\hat{\Phi}$ and $\hat{\Psi}$ can be defined as below:

$$\hat{\Gamma} = \begin{bmatrix} \boldsymbol{C} \\ \boldsymbol{C\hat{A}} \\ \vdots \\ \boldsymbol{C\hat{A}^{f-1}} \end{bmatrix} \quad \text{B.74}$$

$$\hat{\Phi} = \begin{bmatrix} \boldsymbol{D} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{C\hat{B}} & \boldsymbol{D} & & \vdots \\ \vdots & & \ddots & \boldsymbol{0} \\ \boldsymbol{C\hat{A}^{f-2}\hat{B}} & & & \boldsymbol{C\hat{B}\ D} \end{bmatrix} \quad \text{B.75}$$

$$\hat{\Psi} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ \boldsymbol{CK} & 0 & & \vdots \\ \vdots & & \ddots & 0 \\ \boldsymbol{C\hat{A}^{f-2}K} & & \boldsymbol{CK} & 0 \end{bmatrix} \quad \text{B.76}$$

It is assumed that matrix $\boldsymbol{\hat{A}}$ can be approximated by truncating the ARX model just as it is done with CCA method [75]. The aim is to estimate a high order ARX model first and get estimates of the impulse response coefficients $D, C\hat{A}^k\hat{B}$ and $C\hat{A}^k K$ for $k = 0,1,2,3,\ldots,f-2$. The equation B.73 can be written as below if the estimated impulse response coefficients to estimate $\hat{\Phi}$ and $\hat{\Psi}$ [75].

$$z(t) \stackrel{\text{def}}{=} y_f(t) - \hat{\Phi} u_f(t) - \hat{\Psi} y_f(t) = \hat{\Gamma}\hat{x}(t) + e_f(t) \quad \text{B.77}$$

$$\hat{x}(t) = \Delta p(t) \quad \text{B.78}$$

In this case Δ is a matrix of unknown coefficients and $p(t)$ a vector containing delayed inputs and outputs of p steps back:

$$p(t) = [y^T(t-1)\, y^T(t-2) \ldots y^T(t-T)\, u^T(t-1)\ u^T(t-2) \ldots u^T(t-p)] \quad \text{B.79}$$

Just as in CCA, the equation 2.77 is regarded as a low linear regression in r̂Δ[75]. But for CCA the estimation of Δ is done so that the state estimation of can subsequently be done. This can be done by performing a correlation analysis equation on equation 2.77and the equation below follows:

$$M = (R_{zz})^{-\frac{1}{2}} \ (R_{zp})(R_{pp})^{-\frac{1}{2}} \quad \text{B.80}$$

And the sample correlation matrix between two signals $z(t)$and $p(t)$ can be defined as below;

$$R_{zp} = \frac{1}{N}\sum_{t=1}^{N} z(t)p^T(N-t) \quad \text{B.81}$$

The computation of the singular value decomposition of $M = USV^T$, as the next step in CCA, where the CCA estimate of Δ can be written as;

$$\hat{\Delta} = \ V_n^T(R_{pp})^{-\frac{1}{2}} \quad \text{B.82}$$

And the state sequence becomes;

$$\hat{x}(t) = \ V_n^T(R_{pp})^{-\frac{1}{2}} p(t) \quad \text{B.83}$$

Using the linear regression the system matrices can be obtained in the state space model equations by replacing the true state with the estimated state (that is from equation B.71 and equation B.83 [75]).

This method that was proposed by Jansson is better in its performance than many subspace identification methods, CCA, N4SID and MOESP [75]. This SSARX subspace identification method can be applied to both open loop and closed loop data.

B.6.1.4 Virtual Closed loop Algorithm

This is one of the subspace system identification methods which is based on indirect approach and was proposed by Agüero and Goodwin [76]. This method requires the use of the true controller since is one of the indirect procedures, which causes the problems when the true controller is non linear. The true controller, in case of the MPC, may have a high gain in critical areas and this mat mask the plant response. The development of this identification method is based on virtual feedback by using a known linear virtual controller [76] in the analysis, irrespective of the true typical non-linear controller. It is assumed that a linear controller is known and that it will stabilize the process and can be defined as below:

$$\bar{\boldsymbol{C}} = \boldsymbol{P}/\boldsymbol{L} \quad \text{B.84}$$

The introduction of the observer polynomial is done, where the virtual closed loop is defined as $N = E - L$.The virtual loop construction can be illustrated in figure below. The virtual controller added and subtracted not to modify the true closed loop but is used to estimate the system$G$, if the system is in open loop unstable or marginally stable and a subspace method is used [76]. The virtual reference for the system is as below:

$$\bar{u}_t = \frac{L}{E}u_t + \frac{P}{E}y_t = u_t - \frac{N}{E}u_t + \frac{P}{E}y_t \qquad \text{B.85}$$

From the equation above and the system in figure 2.3 one can come up with this relationship [76]:

$$\begin{bmatrix} y_t \\ u_t \end{bmatrix} = \begin{bmatrix} \frac{B_0E}{A_0L+B_0P} \\ \frac{A_0E}{A_0L+B_0P} \end{bmatrix} \bar{u}_t + \begin{bmatrix} \frac{A_0L}{A_0L+B_0P} \\ \frac{A_0P}{A_0L+B_0P} \end{bmatrix} v_t = \bar{T}_0\bar{u}_t + \bar{H}_0^{vcl}v_t \qquad \text{B.86}$$

Where the process G is defined as$G = \frac{B_0}{A_0}$, and $\bar{T}_0 = \begin{bmatrix} T_0^y & T_0^u \end{bmatrix}^T$ is a virtual closed loop function and $\bar{H}_0^{vcl} = \begin{bmatrix} \bar{H}_0^y & \bar{H}_0^u \end{bmatrix}^T$ is the noise transfer [76].

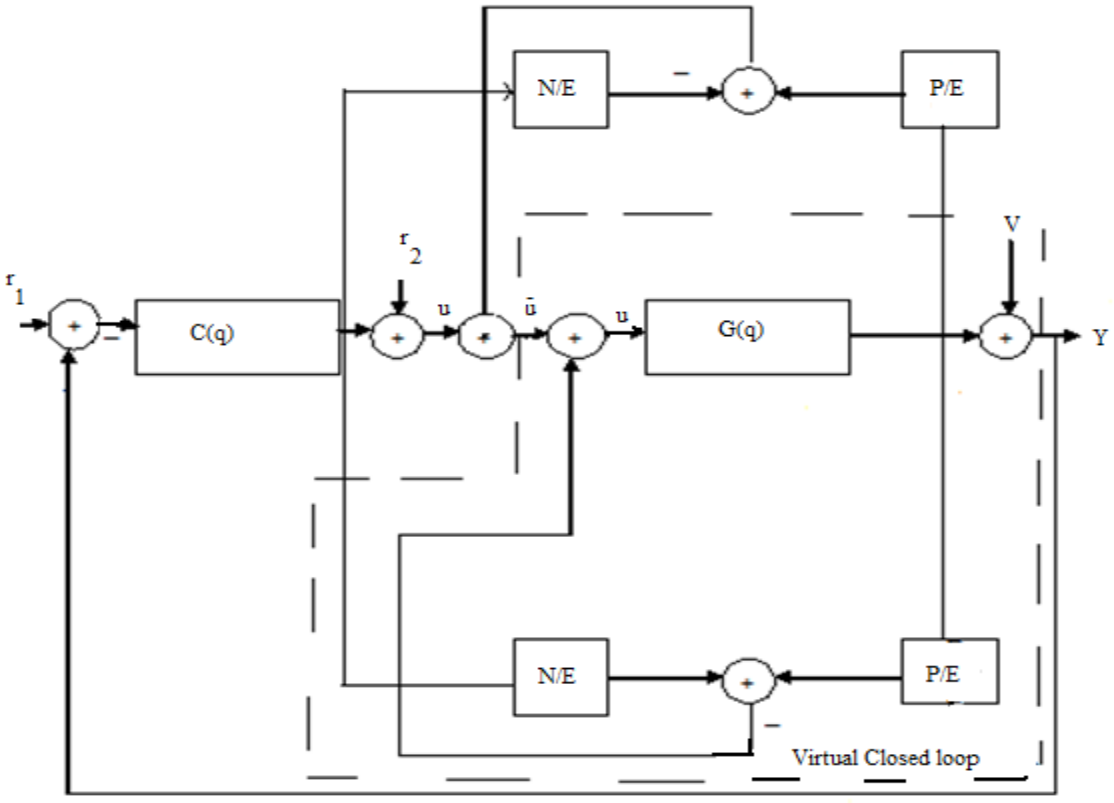

Figure B.4: Virtual Closed loop construction.

L and E are the polynomials and the observer polynomials respectively. It is also stated in [76] that a consistent estimate of $\overline{\boldsymbol{T}}_{\boldsymbol{0}}$ can be obtained by executing three steps that are;

The calculation of a matrix that contains the d-step ahead predictions.

The calculation of the states using the singular value decomposition.

The state-space matrices are estimated as linear regression.

The process estimate can be calculated using the direct approach from either $\widehat{\boldsymbol{T}}_{\boldsymbol{0}}^{\boldsymbol{y}}$ or $\widehat{\boldsymbol{T}}_{\boldsymbol{0}}^{\boldsymbol{u}}$ estimates as below;

$$\hat{G} = \frac{\widehat{\boldsymbol{T}}^{\boldsymbol{y}}\bar{\boldsymbol{L}}}{1-\widehat{\boldsymbol{T}}^{\boldsymbol{y}}\bar{\boldsymbol{P}}} \text{Or} \hat{G} = \frac{\widehat{\boldsymbol{T}}^{\boldsymbol{u}}\bar{\boldsymbol{L}}}{1-\widehat{\boldsymbol{T}}^{\boldsymbol{u}}\bar{\boldsymbol{P}}} \quad \text{B.87}$$

There is an alternative to recover the process estimates, thus by using the joint input-output approach and then $\hat{G}$ becomes:

$$\hat{G} = \frac{\widehat{\boldsymbol{T}}^{\boldsymbol{y}}}{\widehat{\boldsymbol{T}}^{\boldsymbol{u}}} \quad \text{B.88}$$

This allows the identification of accurate models with long prediction horizons operating in closed-loop. The implementation virtual closed loop helps to avoid the restrictions that are associated with indirect approach. The true controller is not needed to be known or used in the identification process. This allows the use of subspace methods and avoiding the numerical problems.

B.6.1.5 Two Stages ORT-based Subspace Method

It was proposed in [77] that this method can be used for identifying open-loop systems operating in closed loop. This method is based on two successive orthogonal decompositions (ORT). The first is the LQ-decomposition which is used for data processing and calculates the deterministic components of the input-output process. The second stage is the decomposition that uses the ORT method to compute the system matrices.

This method is another version of the two stage orthogonal decomposition subspace (TSODS), which is a two-step projection method.

The extended state input-output matrix after the LQ-decomposition can be written as:

$$\hat{Y}_f = \mathrm{r}_k \hat{X}_k + \psi_k \hat{U}_f \quad \text{B.89}$$

Where the matrices r_k and ψ_k are the extended observability matrix and Toeplitz matrix. The ORT method to identify the state-space matrices $\begin{bmatrix} A_{pd} & B_{pd} \\ C_{pd} & D_{pd} \end{bmatrix}$ can be described as $R, U\ and\ Y$ [8] in the Hankel matrices, as follows:

Computation of the LQ-decomposition.

Computation of the LQ-decomposition of the deterministic components.

Estimation of the extended observability matrix by singular value decomposition.

Computation of the system matrices using the self-invariance property of the extended observability matrix$[C_{pd} \quad A_{pd}]$.

The term ψ_k $[B_{pd} \quad D_{pd}]$ is linear with respect to the parameters, such that $[B_{pd} \quad D_{pd}]$ can easily be found.

The advantage of this method which was proposed by Katayama and Tanaka [77] is that it removes stochastic components in the data by pre-processing. The advantage of the two-stage ORT is that it can easily be applied to multivariable systems through the direct approach.

Discussions

A number of the most used subspace identification methods for both open and closed systems have been discussed in this section. N4SID, CCA, MOESP and IV-4SID are usually used in open loop system hence they are sometimes called conventional open loop subspace algorithms. This is so because it is not easy subspace system identification on closed loop process data due to an extended future horizon which introduces correlation between the past input data and the future noise [53]. The added non causal terms in the system model used for subspace projection leads to inflated variances [74]. To use the close loop identification methods it is required that the inputs must be uncorrelated with the noise, which poses a problem for one to use these methods.

The two stages ORT- based subspace method use the ORT projection to filter out the stochastic noise components [77]. The PARSIM-E method and the SSARX method address the problem of the required condition of no correlation of the past input data and the future noise by the use of the pre-estimation which separates the two [74]. The Virtual feedback eliminates the requirement of a known controller model [77], which makes it possible to use the indirect approach that requires the use of the a-prior controller for information [54].

BIBLIOGRAPHY AND REFERENCES

[1]. E.W. Jacobsen. *Studies on dynamics and Control of Distillation Columns.* PhD thesis, University of Trondheim, Norway, 1991.

[2]. Y.Q. Cheng, Y.M.Zhuang, and J.Y Yang. *Optimal Fisher Discriminant analysis using the rank decomposition. Pattern Recognition,* 25: 101 – 111, 1992.

[3]. R.A.Fisher. *The use of multiple measurements in taxomic problems,* Ann Eugenics, 7:179 – 188, 1936.

[4]. S. Wold. *Cross – validatory estimation of components in factor and principal components models.* Technometrics, 20:397 – 405, 1978.

[5]. J.E.Jackson and G.S. Mudholkar. Control Procedures for residuals associated with principal component analysis. Technometrics, 21:341 – 349, 1979.

[6]. T.Kourti and J.F MacGregor. Multivariate SPC methods for process and product monitoring, Journal of quality technology, 28:409 – 428, 1996.

[7]. P. Miller, R.E. Swanson and C.E.Heckler. *Contribution plots: A missing link in multivariate quality control,* Applied mathematics and computer science, 8:775 – 792, 1998.

[8]. B.M. Wise, N.L. Ricker and D.F.Veltkamp. *Upset and sensor failure detection multivariate processes*. Technical report Eigenvector research, Manson, Washington, 1989.

[9]. L.H. Chiang, E.L. Russell and R.D. Braatz. *Fault detection and diagnosis in industrial systems*. Springer, 2001.

[10]. R.Dunia, S.J. Qin, T.F. Edgar and T.J. McAvoy. *Identification of faulty sensors using principal component analysis*. AIChe J., 42: 2797 – 2812, 1996.

[11]. A. Negiz and A. Cinar. *On the detection of multiple sensor abnormalities in multivariate processes*. In proc. Of the American conference, pg. 2364 – 2368, Piscataway, New Jersey, 1992. IEEE Press.

[12]. D.M. Himes, R.H. Storer and C. Georgakis. *Determination of the number of principal components for disturbance detection and isolation*. In proc. Of American control conference, pages 1279 – 1283, Piscataway, New Jersey, 1994. IEEE Press.

[13]. J. Zhang, E. Martin and A.J.Morris. *Fault detection and classification through multivariate statistical techniques*. In proc. Of the American control conference, pages 751 – 755, Piscataway, New Jersey, 1995. IEEE Press.

[14]. A.C.Raich and A. Cinar. *Statistical process monitoring and disturbance isolation in multivariate continuous processes*. In proc of the IFAC Conf on advanced control of chemical processes, pages 427 – 435, New York, 1994. Pergamon Press.

[15]. A.C. Raich and A. Cinar. *Multivariate statistical methods for monitoring continuous processes: Assessment of discriminatory power disturbance and models and diagnosis of multiple disturbances*. Chemometrics and intelligent laboratory systems, 30:37 – 48, 1995.

[16]. A.C. Raich and A. Cinar. *Statistical process monitoring and disturbance diagnosis in multivariate continuous processes.* AIChe J., 42:995 – 1009, 1996.

[17]. W.J. Krzanowski. *Between – group comparison of principal components*. J. Amer. Stat.Assn, 74:703 – 706, 1979.

[18]. M.A. Kramer. *Non linear principal component analysis using auto associative neural networks.* AIChe J., 37:233 – 243, 1991.

[19]. W.W.S. Wei. *Time series analysis*. Addison – Wesley, Reading, Masschusetts, 1994.

[20]. B.M. Wise and N.B. Gallanger. *The process chemometrics approach to process monitoring and fault detection.* J. of process control, 6:329 – 348, 1996.

[21]. W.Ku, R.H. Storer and C. Georgakis. *Disturbance detection and isolation by dynamic principal component analysis*. Chemometrics and intelligent laboratory systems, 30:179 – 196, 1995.

[22]. A. Negiz and A. Cinar. *Statistical monitoring of multivariable dynamic processes with state space models.* AIChE J., 43: 2002 – 2020, 1997.

[23]. D.Dong and T.J. McAvoy. *Non linear principal component analysis: Based on principal curves and neural networks.* Computers and chemical engineering, 20: 65 – 78, 1996.

[24]. G. Chen and T.J. McAvoy. *Predictive on line monitoring of continuous processes*. J. of process control, 8:409 – 420, 1997.

[25]. C. Bishop. *Neural networks for pattern recognition*. Clarendon press, New York, 1995.

[26]. R.Dunia, S.J.Qin and T.F.Edgar. *Multivariable process monitoring using non linear approaches*. In proc. Of the American control conf., pages 756 – 760, Pitscataway, New Jersey, 1995. IEEE Press.

[27]. P.Nomikos and J.F. MacGregor. *Monitoring batch processes using multiway principal component analysis*. AIChE J., 40: 1361 – 1375, 1994.

[28]. R.B. Crosier. *Multivariate generalizations of cumulative sum quality control schemes*. Technometrics, 30:291 – 303, 1998.

[29]. C.A. Lowry and W.H Woodall. *A multivariate exponentially weighted moving average control chart*. Technometrics, 34: 46 – 53, 1992.

[30]. J.F. MacGregor. *Statistical process control of multivariate processes*. In proc.of the IFAC int.symp. On advanced control chemical processes, pages 427 – 435, New York, 1994. Pergamon press.

[31]. J.J. Pignatiello, Jr and G.C. Runger. *Comparisons of multivariate CUSUM charts*. J. of quality Technology, 22: 173 – 186, 1990.

[32]. W.H.Woodall and M.M. Ncube. *Multivariate CUSUM quality control procedures*. Technometrics, 27: 285 – 292, 1985.

[33]. J.D. Healy. *A note of multivariate CUSUM procedures*. Technometrics, 29: 409 – 412, 1987.

[34]. J.V. Cresta, T.E. Marlin and J.F MacGregor. *Multivariable statistical monitoring of process operating performance*. Can. J. of chem. Eng., 69:35 – 47, 1991.

[35]. M.J. Piovoso, K.A.Kosanovich and R.K.Pearson. *Monitoring process performance in real time*. In proc. Of the American control conference, pages 2359 – 2363, Piscataway, New Jersey, 1992. IEEE Press.

[36]. W.W. Hines and D.C. Montgomery. *Probability and statistics in engineering and management science*. John Wiley and sons, New York, 3rd edition, 1990.

[37]. B.M. Wise, N.L. Ricker, D.J. Velkamp and B.R. Kowalski. A theoretical basis for the use of principal component analysis monitoring multivariate processes. Process control and quality, 1:41 – 51, 1990.

[38]. J.E. Jackson. *A users guide to principal components*. John Wiley& sons, New York, 1991.

[39]. S.Valle, W.Li and S.J. Qin. Selection of the number of components: The variance of the reconstruction error criterion with a comparison to other methods. Ind.Eng. Chem. Res. Vol 38, 4389 – 4401, 1999.

[40]. S.Wold, K.Esbensen and P. Geladi. *Principal components analysis.* Chemometrics and intelligent laboratory systems, 2: 37, 1987.

[41]. J.E.Jackson. *Quality control methods for several related variables*. Technometrics, vol 1, 359 – 377, 1959.

[42]. R.A.Johnson and D.W. Wichern. *Applied multivariate statistical analysis*. Prentice hall, New Jersey, 3rd edition, 1992.

[43]. E. Malinowski. Statistical F – Test for abstract factor analysis and target testing. J. chemometrics, 3:46 – 50, 1989.

[44]. O.Nelles, Non linear System identification. Springer, 2001.

[45]. U.Forsell and L.Ljung, "closed loop identification revisited," *Automatica*, Vol.35, pp.1215 -1241, 1999.

[46]. L.Ljung, System identification: Theory for the user. Prentice Hall, 1999.

[47]. H. Zeiger and A.McEwen, "Approximate linear realizations of given dimensions via ho's algorithm", *IEEE transaction on automatic control,* vol.19, no.153, 1974.

[48]. D. Brillinger, *Time series: Data analysis and theory.* San Fransisco, CA: Holden – Day, 1981.

[49]. Y.Zhua and F.Butoyi, "Case studies on closed loop identification for MPC, "*Control engineering for practice,* vol.10, pp.403 – 417, 2002.

[50]. C. Koung and J. MacGregor, "Design identification experiments for robust control, a geometric approach for bivariate processes," *Industrial Engineering and chemical research,* vol.32, pp. 1658 – 1666, 1993.

[51]. L.B Palma, F.J Coito and R.N. Silva "Adaptive Observer Based on Fault Diagnosis Approach Applied to Thermal Plant."*Proceedings of the 10th Mediterranean Conference on Control and Automation*, 2002.

[52]. S.J.Qin, "An overview of subspace identification", *Computers and chemical Engineering,* vol.30, pp. 1502 – 1513, 2006.

[53]. M.Gevers, L.Ljung and P.V. denhof, "Asymptotic variance expressions for closed loop identification and their relevance in identification for control", *selected topics in identification,* vol.9, pp.9 -15, 1996.

[54]. A.Esmaili, J.F. MacGregor and P.A Taylor, "Direct and two-step methods for closed loop identification: a comparison of asymptotic and finite data set performance ", *Journal Process control,* vol.10, pp. 525 – 537, 2000.

[55]. R.Iserman and P. Balle, "Trends in the application of model-based fault detection and diagnosis and technical processes", *Control Engineering practice,* vol.5, no.5, pp.709 -719, 1998.

[56]. L.Ljung, "Asymptotic variance expression for identified black box transfer function model", *IEEE Trans. Auto. Cont.* Vol.30, pp. 834 – 844, 1985.

[57]. Y.Zhu, "Multivariable process identification for mpc: the asymptotic method and its application", *Journal of process control,* vol.8, pp. 101 – 115, April, 1998.

[58]. W.C.Yu and N.Y.Shih," Bi-loop recursive least squares algorithm with forgetting factors", IEEE signal processing letters, vol.13, pp. 505 – 508, August 2008.

[59]. Z.Chao, H. Hua-sheng, B.Wei-min and Z.Luo-ping, "Robust recursive estimation of auto-regressive updating model parameters for real-time flood forecasting", *Journal of hydrology,* vol.349, pp. 376 - 382, 2008.

[60]. I.Eker, "Open-loop and closed loop experimental online identification of a three mass electromechanical system', *Mechatronic,* vol.14, pp. 549 – 565, 2004.

[61]. C. Alexander and R.Trahan, "A comparison of traditional and adaptive control strategies for system with time delay", *ISA Trans,* vol. 40, pp.353 – 368, 2001.

[62]. J.Zhou, "Classical theory of errors and robust estimation "*Aota Geodetica et cartographical sinica,* pp. 115 – 120, 1989.

[63]. A. Ghaffari, J.Roshanian and Tayefi, "Time varying transfer function extraction of an unstable launch vehicle via closed loop identification", *Aerospace Science and technology,* vol.11, pp.238 – 244, 2007.

[64]. T.Soderstrom and P. Stoica, *System identification,* Series in System and control Engineering, NY, USA: Prentice Hall, 1989.

[65]. N.Araki, M.Okada and Y.Korishi, " Parameter identification and swing up control of an acrobat system", *IEEE Transactions Automatica Control,* pp.1040 – 1045, 2005

[66]. S.Julier, J.Ulhlmann and H.F. Durant – Whyte, "A new method for the nonlinear transformation of means and covariances in filters and estimators, "*IEEE Transactions of Automatica Control,* vol.45, no.3, pp.477 – 482, 2000.

[67]. W. Favoreel, B. De Moor and P.Van Overschee, "The subspace system identification for industrial processes", *Journal of process control,* vol.10, pp. 149 – 155, 2000.

[68]. E. De Klerk and J. Craig, " Multivariable closed loop system identification of plants under model predictive control", *in proc of the 13th IFAC symposium on system identification* (P.V. den [68]. H.B. Wahlberg, The Nedellands), pp.411 – 419, IFAC, Elsivier, August 2003.

[69]. P.V. Overchee and B.D.Moor, "N4sid: Subspace algorithms for the identification method of combined deterministic and stochastic systems. *Automatica,* vol.30, no.1, pp. 75 – 93, 1994.

[70]. M.Viberg, "Subspace methods in system identification", *Proc. Of the 10th IFAC symposium on system identification,* vol.1, no.94, pp.1-12, 1994.

[71]. P. Mathieu and M. Muhammad, " Closed loop identification method using a subspace approach", *in a proceedings volume from the 13th IFAC symposium on system identification* vol.1, (Rotterdam, The Nederlands), pp.423 – 428, August 2003.

[72]. S.Qin and L.Ljung, "Parallel Implementation of subspace identification with parsimonious model". IFAC Symposium on system identification 2003.

[73]. S.Qin and L.Ljung, " Closed loop subspace identification with innovation estimation", in a proceedings *Volume from the 13th IFAC symposium of system identification*, vol 2, pp. 861 – 866, August, 2003.

[74]. M. Jansson, "Subspace identification and ARX – Modelling" in a proceedings *Volume from the 13th IFAC symposium of system identification,* pp. 1585 – 1590, August 2003.

[75]. J.C.Agüero and G.C Goodwin, "Virtual closed loop identification: A subspace approach", (Atlantis, Paradise Island, Bahamas), pp. 14 -17, December, 2004.

[76]. T.Katayama and H.tanaka, "An approach to closed loop subspace identification by orthogonal decomposition", *Automatica,* vol.43, 2007.

[77]. S.Skogestad and I.Postethwaile, *Multivariate feedback control: Analysis and Design.* Wiley, 2nd Ed, 2005.

[78]. Sergio Valle, Weihua Li and S. Joe Qin, *Selection Of the number of principal components*, Industrial Engineering Chemical Res, vol 38, pp. 4339 – 4401, 1999.

[79]. Akaike, H. Information theory and extension of the maximum likelihood principle. In proceedings 2nd international symposium on information theory, Petrov and Caski, 1994; pp 267 – 281.

[80]. Rissanen.J. Modelling by shortest data description .Automatica, 1978, vol14, pp.4655 – 471.

[81]. M.Wax, T. Kailath. Detection of signals by information criteria. *IEEE Trans. Acoust.Speech Signal process.* ASSP- vol33, 1985, 387 – 392.

[82]. E.R. Malinowski. Determination of the number of factors and the experimental error in a data matrix. *Anal.Chem.* 1977, vol 49(4), pp. 612 -617.

[83]. E.R. Malinowski. *Factor Analysis in chemistry;* Wiley – Interscience: New York, 1991.

[84]. R.B. Cattell. The scree Test for the number of factors. *Multivariate Behaviour. Res,* pp. 245 – 276, April 1966.

[85]. S.Wold. Cross Validatory estimation of the number of components in factor and principal component analysis. *Technometrics,* 1978, vol 20, pp.397 – 406.

[86]. M.Tamura and S. Tsujita, A study on the principal components and sensitivity of fault detection using PCA. Computers and Chemical Engineering, vol 31, pp. 1035 – 1046, 2007.

[87]. R.Dunia and S.J.Qin, Joint diagnosis of process and sensor faults using principal component analysis. *Control Engineering Practice*, vol.6, pp.245 -255, 1998.

[88]. J.E.Jackson and Mudhokar, Control Procedures for residuals associated with principal component analysis. *Technometrics*, vol.21, pp. 341 – 349, 1979.

[89]. R.S LaFleur and J.A. Taylor, Control of a Thermal System, AE/ME 401- Spring, 2005.

[90]. S.J.Qin. Determination of the number of principal components for the best reconstruction. Journal of process control vol 10 pp.25 -28, 2000.

[91].Y. Chen and L. Lan. Energy and Buildings, vol 41, pp. 881–887, 2009.

[92]. F.Xiao, S. Wang, X. Xu, 'An isolation enhanced PCA with expert based multivariate decoupling for sensor FDD in air conditioning systems, 'Applied Thermal Engineering', 2009.

[93]. S.Wang, Q. Zhou and F. Xiao, A system level detection and diagnosis strategy for HVAC systems involving sensor faults, 'Energy and building', vol 4, pp. 477 – 490, 2010.

[94]. David Zumoffa and M. Basualdo, A systematic approach for design of optimal monitoring for a large scale processes. 'Industrial chemical engineering', vol 49, pp 1749 – 1761, 2010.

[95]. D.Ming – Cong, I.Akira and E. Kazunori, Fault Detection system design for actuator of a thermal process using operator based, 'Automatica Sinica, vol 36, April 2010.

[96]. M.Braae, MIMO Control System Design Lecture notes, University of Cape Town, 2006.

[97]. S. Suwed, Control Monitoring, Testing and Measurement, August 2010.

[98]. Donaldson, Circuit wiring diagram, Volkswagen- Golf heater, July 2010.

[99]. M. Tseou, Proton Beam Steering, University of Cape Town, 2008

[100]. B. Patel, Investigation into Fault Detection and Diagnosis Techniques, University of Cape Town, April 2000.

[101]. R. Isermann," Fault Detection and Management in Industrial Plants", 2005

[102]. J. Gertler, Survey of Model based Failure Detection and Isolation in Complex Plant IEEE Control Systems Magazine, 1988.

[103]. Iserman and P.Balle, Trends in Application of Model Based Fault Detection and Diagnosis of Technical Processes. IFAC Symposium SAFE PROCESS'96, Paper 7f-011, San Fransisco, 1996.

[104]. P.M. Frank, Fault Diagnosis in Dynamic Systems Using Analytical and Knowledge Based Redundancy – a survey and some new results. Automatica volume 26, number 3, PP. 459 – 474, 1990.

[105]. E.Y.Chow and A.S Wilsky, Analytical Redundancy and the Design of Robust Detection Systems. IEEE Trans – Automatic Control, AC – 29, PP 603 – 614, 1984.

[106]. T.R. Patton and P.Miller, Robust Fault Diagnosis of two Pump System – Case Study. IFAC Symposium SAFEPROCESS'96, Paper Tf – 031, San Fransisco, USA, 1996.

[107]. J. Liu and F. Gao, ”A frequency domain step domain response identification method for continuous time delay”. Journal of process control, 2010.

[108]. Lee and Edgar, “Simple graphical method for noisy pulse and step response”. Chemical engineering science, 2010.